

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

NOVAS METODOLOGIAS ANALÍTICAS NA AUTENTICAÇÃO DE  
SUCOS DE UVA INTEGRAL ORGÂNICOS EMPREGANDO DADOS  
DE TÉCNICAS DE ESPECTROMETRIA MOLECULAR E  
FERRAMENTAS QUIMIOMÉTRICAS

CARLOS HENRIQUE JUNGES

ORIENTADOR: PROF. DR. MARCO FLÔRES FERRÃO  
COORIENTADOR: DR. CELITO CRIVELLARO GUERRA

TESE DE DOUTORADO

Porto Alegre, Abril de 2023

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

CARLOS HENRIQUE JUNGES

**NOVAS METODOLOGIAS ANALÍTICAS NA AUTENTICAÇÃO  
DE SUCOS DE UVA INTEGRAL ORGÂNICOS EMPREGANDO  
DADOS DE TÉCNICAS DE ESPECTROMETRIA MOLECULAR E  
FERRAMENTAS QUIMIOMÉTRICAS**

Tese apresentada como requisito parcial para a  
obtenção do grau de Doutor em Química

Prof. Dr. Marco Flôres Ferrão  
Orientador

Dr. Celito Crivellaro Guerra  
Coorientador

Porto Alegre, Abril de 2023

## DECLARAÇÃO DE AUTORIA E ORIENTAÇÃO

A presente tese foi realizada inteiramente elaborado pelo autor, exceto as colaborações as quais serão devidamente citadas nos agradecimentos, no período entre Outubro/2019 e Abril/2023, no Instituto de Química da Universidade Federal do Rio Grande do Sul sob Orientação do Professor Doutor Marco Flôres Ferrão e Coorientação do Doutor Celito Crivellaro Guerra. A tese foi julgada adequada para a obtenção do título de Doutor em Química pela seguinte banca examinadora:

### Comissão Examinadora:

---

Prof. Dr. Adilson Ben da Costa  
UNISC

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Silvana Mariela Azcarate  
UNLPam

---

Prof. Dr. João Henrique Zimnoch dos Santos  
UFRGS

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Carla Sirtori  
UFRGS

---

Prof. Dr. Marco Flôres Ferrão  
Orientador - UFRGS

---

Dr. Celito Crivellaro Guerra  
Coorientador - EMBRAPA

---

Carlos Henrique Junges  
Aluno - UFRGS

## **DEDICATÓRIA**

Ao meu pai Carlos Alberto Junges (*in memoriam*).

## EPÍGRAFE

“O segredo da criatividade está em dormir bem e abrir a mente às possibilidades infinitas. O que é um homem sem sonhos?”

ALBERT EINSTEIN

## AGRADECIMENTOS

À minha mãe e irmã, Rejane e Paula, pelo carinho, suporte, compreensão e incentivos concebidos no decorrer deste período de aprendizagem.

Ao meu querido orientador, Prof. Dr. Marco Flôres Ferrão, pela amizade, ensinamentos e apoio durante o desenvolvimento deste trabalho. Ferrão, agradeço imensamente pela oportunidade de trabalharmos juntos, pela confiança depositada em mim e por todo conhecimento transmitido.

Ao meu coorientador, Dr. Celito Crivellaro Guerra, pela receptividade, colaborações e sugestões.

Ao Prof. Dr. Adriano de Araújo Gomes pela amizade e valiosos conhecimentos e auxílios ao longo desta trajetória. Muito obrigado, Adriano.

A todos os meus colegas do Laboratório de Quimiometria e Instrumentação Analítica (LAQIA) pela amizade, acolhimento, ajudas e convivência, em especial para Fernanda Böck, Danielle Kochenborger e Mateus Schneider.

Aos professores das bancas do Exame de Qualificação e da Tese pelas colaborações no aprimoramento deste trabalho.

À Embrapa Unidade Uva e Vinho de Bento Gonçalves-RS pela cessão das amostras de suco de uva, essencial para concepção deste trabalho.

Ao Programa de Pós-Graduação em Química (PPGQ) pela assistência.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro.

À Universidade Federal do Rio Grande do Sul (UFRGS) pela estrutura, equipamentos e fomento à educação e ciência.

# PRODUÇÃO CIENTÍFICA

## 1. Artigos publicados durante o doutorado

1.1 Carlos H. Junges, Celito C. Guerra, Natalia A.P. Canedo-Reis, Adriano A. Gomes, Fernando S. Diogo, Marco F. Ferrão. Grape juice classification with respect agricultural production system by means of visible spectroscopy chemometrics assisted. *Journal of Food Composition and Analysis*, 114, 104793, 2022.

1.2 Carlos H. Junges, Celito C. Guerra, Adriano A. Gomes, Marco F. Ferrão. Green analytical methodology for grape juice classification using FTIR spectroscopy combined with chemometrics. *Talanta Open*, 6, 100168, 2022.

1.3 Carlos H. Junges, Celito C. Guerra, Natalia A.P. Canedo-Reis, Adriano A. Gomes, Marco F. Ferrão. Discrimination of whole grape juice using fluorescence spectroscopy data with linear discriminant analysis coupled genetic and ant colony optimisation algorithms. *Analytical Methods*, 15, 187-195, 2023.

## 2. Artigos publicados em colaboração

2.1 Natalia A.P. Canedo-Reis, Celito C. Guerra, Leticia F. da Silva, Luísa C. Wetzstein, Carlos H. Junges, Marco F. Ferrão, Ana M. Bergold. Fast quantitative determination of phenolic compounds in grape juice by UPLC-MS: method validation and characterization of juices produced with different grape varieties. *Journal of Food Measurement and Characterization*, 15, 1044-1056, 2021.

2.2 Natalia A.P. Canedo-Reis, Flávia Suelen, Daiana S. Ávila, Celito C. Guerra, Letícia F. da Silva, Carlos H. Junges, Marco F. Ferrão, Ana M. Bergold. Grape juice reduces the effects of amyloid  $\beta$  aggregation phenotype and extends the longevity in *Caenorhabditis elegans*. *Nutritional Neuroscience*, 2022. (<https://doi.org/10.1080/1028415X.2022.2140394>).

### **3. Trabalhos apresentados em eventos**

3.1 Carlos H. Junges, Fernanda C. Böck, Marco F. Ferrão, Celito C. Guerra. pHmetro por imagen via extracto de frijol negro y PhotoMetrix UVC®. 10° Congreso Argentino de Quimica Analítica, Setembro, 2019, Santa Rosa, La Pampa, Argentina.

3.2 Natalia A.P. Canedo-Reis, Carlos H. Junges, Ana M. Bergold, Marco F. Ferrão, Celito C. Guerra. Caracterização de sucos de uva produzidos no sul do Brasil empregando análise exploratória pelos métodos de HCA e de PCA. XI Workshop de Quimiometria, Novembro, 2020, Campina Grande, Paraíba, RS.

3.3 Carlos H. Junges, Celito C. Guerra, Natalia A.P. Canedo-Reis, Adriano A. Gomes, Fernando S. Diogo, Marco F. Ferrão. Grape juice classification with respect agricultural production system by means of visible spectroscopy chemometrics assisted. 20° Encontro Nacional de Química Analítica, Setembro, 2022, Bento Gonçalves, RS.



## SUMÁRIO

<b>LISTA DE FIGURAS</b> .....	11
<b>LISTA DE TABELAS</b> .....	13
<b>LISTA DE ABREVIATURAS</b> .....	14
<b>RESUMO</b> .....	16
<b>ABSTRACT</b> .....	17
<b>1. INTRODUÇÃO</b> .....	18
1.1 OBJETIVO GERAL.....	19
<b>1.1.1 Objetivos Específicos</b> .....	19
<b>2. REVISÃO BIBLIOGRÁFICA</b> .....	21
2.1 BREVE HISTÓRICO DA VITICULTURA MUNDIAL E BRASILEIRA.....	21
2.2 SITUAÇÃO ATUAL DA PRODUÇÃO DE UVAS E SUCOS DE UVA.....	22
2.3 SISTEMAS CONVENCIONAL E ORGÂNICO DE PLANTIO DE UVA.....	27
2.4 CARACTERÍSTICAS DA UVA E DAS VIDEIRAS.....	28
2.5 DEFINIÇÃO E CLASSIFICAÇÃO DO SUCO DE UVA.....	30
2.6 COMPOSIÇÃO QUÍMICA DO SUCO DE UVA.....	31
2.7 PROCESSO DE ELABORAÇÃO DO SUCO DE UVA .....	34
2.8 CONTROLE DE QUALIDADE DOS SUCOS DE UVA .....	37
<b>2.8.1 Sólidos Solúveis</b> .....	38
<b>2.8.2 Sólidos Insolúveis</b> .....	38
<b>2.8.3 Acidez Total</b> .....	39
<b>2.8.4 Acidez Volátil</b> .....	39
<b>2.8.5 Sorbitol, Florizina, Edulcorantes, Álcool Etilico e Corante Artificial</b> .....	39
2.9 TÉCNICAS DE ESPECTROMETRIA MOLECULAR .....	40
<b>2.9.1 Espectrometria na Região do Ultravioleta-Visível</b> .....	40
<b>2.9.2 Espectrometria na Região do Infravermelho</b> .....	45
<b>2.9.3 Espectrometria de Fluorescência Molecular</b> .....	50
2.10 FERRAMENTAS QUIMIOMÉTRICAS.....	52
<b>2.10.1 Técnicas de Classificação Multivariada</b> .....	52
2.10.1.1 Análise Discriminante Linear (LDA).....	53
2.10.1.2 Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA). 55	
<b>2.10.2 Técnicas de Seleção de Variáveis</b> .....	58
2.10.2.1 Algoritmo Genético (GA) .....	59

2.10.2.2	Algoritmo de Otimização por Colônia de Formigas (ACO) .....	61
2.10.2.3	Algoritmo Stepwise (SW) .....	63
2.10.2.4	Algoritmo das Projeções Sucessivas (SPA) .....	65
<b>2.10.3</b>	<b>Pré-Tratamento dos Dados</b> .....	<b>67</b>
<b>3.</b>	<b>METODOLOGIA EXPERIMENTAL</b> .....	<b>69</b>
3.1	REAGENTES E MATERIAIS .....	69
3.2	AMOSTRAS DE SUCOS DE UVA .....	70
3.3	PREPARAÇÃO DAS AMOSTRAS .....	72
3.4	MEDIDAS DE ESPECTROSCOPIA NA REGIÃO DO VISÍVEL .....	73
3.5	MEDIDAS DE ESPECTROSCOPIA NA REGIÃO DO INFRAVERMELHO MÉDIO....	74
3.6	MEDIDAS DE ESPECTROSCOPIA DE FLUORESCÊNCIA .....	74
3.7	ANÁLISE QUIMIOMÉTRICA.....	75
3.8	VALIDAÇÃO ANALÍTICA DOS MÉTODOS .....	75
<b>4.</b>	<b>RESULTADOS E DISCUSSÃO</b> .....	<b>77</b>
4.1	CLASSIFICAÇÃO DE SUCOS DE UVA COM RESPEITO AO SISTEMA DE PRODUÇÃO AGRÍCOLA POR MEIO DE ESPECTROSCOPIA VISÍVEL ASSISTIDA POR QUIMIOMETRIA .....	77
<b>4.1.1</b>	<b>Espectros Visíveis em Diferentes pHs</b> .....	<b>77</b>
<b>4.1.2</b>	<b>Modelos LDA e PLS-DA</b> .....	<b>79</b>
<b>4.1.3</b>	<b>Seleção de Variáveis</b> .....	<b>85</b>
4.2	METODOLOGIA ANALÍTICA VERDE PARA CLASSIFICAÇÃO DE SUCOS DE UVA USANDO ESPECTROSCOPIA FTIR COMBINADA COM QUIMIOMETRIA ..88	
<b>4.2.1</b>	<b>Espectros no Infravermelho Médio</b> .....	<b>88</b>
<b>4.2.2</b>	<b>Modelos LDA e PLS-DA</b> .....	<b>89</b>
<b>4.2.3</b>	<b>Seleção de Variáveis</b> .....	<b>94</b>
4.3	DISCRIMINAÇÃO DE SUCO DE UVA INTEGRAL USANDO DADOS DE ESPECTROSCOPIA DE FLUORESCÊNCIA COM ANÁLISE DISCRIMINANTE LINEAR ACOPLADA AOS ALGORITMOS GENÉTICO E OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS .....	96
<b>4.3.1</b>	<b>Espectros de Fluorescência</b> .....	<b>97</b>
<b>4.3.2</b>	<b>Modelos LDA e PLS-DA</b> .....	<b>99</b>
<b>4.3.3</b>	<b>Seleção de Variáveis</b> .....	<b>102</b>
<b>5.</b>	<b>CONCLUSÕES</b> .....	<b>105</b>
<b>6.</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	<b>107</b>

## LISTA DE FIGURAS

<b>Figura 1.</b> Produção média de uva por continente no período de 2010 a 2020 .....	23
<b>Figura 2.</b> Evolução da quantidade produzida e área de colheita de uva no mundo no período de 1980 a 2020 .....	23
<b>Figura 3.</b> Evolução da quantidade produzida e área de colheita de uva no Brasil no período de 1980 a 2020 .....	24
<b>Figura 4.</b> Quantidade média produzida de uva pelos estados brasileiros no período de 2018 a 2020 .....	24
<b>Figura 5.</b> Quantidade média produzida de uva pelos municípios do RS no período de 2018 a 2020 .....	26
<b>Figura 6.</b> Modelo típico do cultivo de videira em pérgola ou latada.....	29
<b>Figura 7.</b> Fórmula estrutural de alguns componentes químicos presentes no suco de uva..	33
<b>Figura 8.</b> Etapas do processamento industrial para elaboração do suco de uva.....	35
<b>Figura 9.</b> Espectro eletromagnético com os tipos de radiação e os intervalos de comprimentos de onda e de frequência .....	41
<b>Figura 10.</b> Exemplo de espectro na região do ultravioleta-vísivel .....	42
<b>Figura 11.</b> Representação do feixe incidente de luz sobre um fluido em uma cubeta com caminho óptico definido $b$ e do feixe transmitido da radiação de luz .....	43
<b>Figura 12.</b> Representação de um espectrofotômetro de (a) feixe simples e (b) duplo feixe .....	44
<b>Figura 13.</b> Exemplo de espectro na região do infravermelho médio.....	45
<b>Figura 14.</b> Modos de vibrações moleculares de deformação axial e angular.....	48
<b>Figura 15.</b> Representação de um espectrofotômetro de FTIR.....	49
<b>Figura 16.</b> Diagrama Jablonski com as transições de energia do fenômeno da fluorescência.....	51
<b>Figura 17.</b> Classificação de vários objetos antes e depois da implementação da LDA.	53
<b>Figura 18.</b> Aplicação da técnica LDA com a maximização dos eixos componentes para a separação das classes .....	54
<b>Figura 19.</b> Projeção dos dados ortogonalmente pela técnica LDA .....	55
<b>Figura 20.</b> Organização dos dados para a construção do modelo de classificação usando PLS-DA .....	56
<b>Figura 21.</b> Fluxograma com as etapas envolvidas do Algoritmo Genético.....	60

<b>Figura 22.</b> Comportamento sequencial das formigas explicado pela presença de feromônios nos trajetos.....	61
<b>Figura 23.</b> Fluxograma com as etapas envolvidas do Algoritmo Stepwise.....	64
<b>Figura 24.</b> Fluxograma com as etapas envolvidas nos procedimentos metodológicos para cada estudo proposto .....	69
<b>Figura 25.</b> Suquificador integral para a produção de sucos de uva integral.....	71
<b>Figura 26.</b> Etapas envolvidas na preparação das amostras de suco de uva integral.....	73
<b>Figura 27.</b> Perfis espectrais médios visíveis após modulação de pH para todos os pHs estudados (2, 4, 6, 8, 10 e 12).....	78
<b>Figura 28.</b> Perfis espectrais médios visíveis após a modulação do pH das classes convencional e orgânica para (a) todos os pHs (2, 4, 6, 8, 10 e 12) e (b) pHs alcalinos (8, 10 e 12) .....	79
<b>Figura 29.</b> Gráficos de função discriminante das amostras de treinamento e teste dos melhores modelos por GA-LDA para (a) todos os pHs (2, 4, 6, 8, 10 e 12) e (b) pHs alcalinos (8, 10 e 12).....	84
<b>Figura 30.</b> Alterações estruturais de uma antocianina genérica em meio aquoso em função da variação do pH .....	85
<b>Figura 31.</b> Seleção das variáveis mais relevantes dos melhores modelos por GA-LDA para (a) todos os pHs (2, 4, 6, 8, 10 e 12) e (b) pHs alcalinos (8, 10 e 12) .....	87
<b>Figura 32.</b> Perfis espectrais FTIR dos valores médios das classes convencional e orgânica .88	
<b>Figura 33.</b> Gráficos de função discriminante das amostras de treinamento e teste dos melhores modelos por (a) GA-LDA, (b) SW-LDA, (c) ACO-LDA e (d) SPA-LDA....	93
<b>Figura 34.</b> Discriminantes de Fisher com seleção dos números de onda ( $\text{cm}^{-1}$ ) mais relevantes e o perfil espectral com as variáveis selecionadas para as classes convencional (regiões laranjas) e orgânica (regiões verdes) dos melhores modelos por GA-LDA (a) e (b), SW-LDA (c) e (d), ACO-LDA (e) e (f), e SPA-LDA (g) e (h), respectivamente .....	95
<b>Figura 35.</b> Matriz excitação-emissão (EEM) de fluorescência para uma amostra de suco de uva integral.....	97
<b>Figura 36.</b> Perfis de emissão para os 10 comprimentos de onda de excitação de fluorescência após (a) desdobramento, (b) remoção dos espalhamentos Rayleigh e Raman, (c) normalização e (d) valores médios dos sinais por classe convencional ou orgânica.....	98
<b>Figura 37.</b> Gráficos de função discriminante das amostras de treinamento e teste dos melhores modelos por (a) ACO-LDA e (b) GA-LDA .....	102
<b>Figura 38.</b> Seleção das variáveis mais relevantes dos melhores modelos por (a) ACO-LDA e (b) GA-LDA .....	103

## LISTA DE TABELAS

<b>Tabela 1.</b> Área plantada e quantidade produzida de uva no Brasil e no Rio Grande do Sul no período de 2008 a 2022 .....	25
<b>Tabela 2.</b> Evolução da produção de sucos de uva nas safras 2017 a 2021 no RS .....	27
<b>Tabela 3.</b> Classificação dos sucos de uva processados segundo a legislação brasileira...	31
<b>Tabela 4.</b> Valores nutricionais do suco de uva .....	32
<b>Tabela 5.</b> Parâmetros físico-químicos avaliados no controle de qualidade dos sucos de uva.....	38
<b>Tabela 6.</b> Regiões espectrais no infravermelho .....	46
<b>Tabela 7.</b> Frequências de absorção no infravermelho de diversos grupos funcionais...	46
<b>Tabela 8.</b> Figuras de mérito dos melhores modelos de classificação para dados de espectroscopia na região do visível .....	81
<b>Tabela 9.</b> Figuras de mérito dos modelos PLS-DA para cada pH individual para dados de espectroscopia na região do visível .....	82
<b>Tabela 10.</b> Figuras de mérito dos melhores modelos de classificação para dados de espectroscopia na região do infravermelho médio .....	89
<b>Tabela 11.</b> Figuras de mérito dos melhores modelos de classificação para dados de espectroscopia de fluorescência.....	99

## LISTA DE ABREVIATURAS

<b>ACO</b>	Algoritmo de Otimização por Colônia de Formigas
<b>ASCII</b>	Código Padrão Americano para o Intercâmbio de Informação
<b>ASP</b>	Páginas de Servidor Ativas
<b>GBIP</b>	Barramento de Interface de Uso Geral
<b>CAPES</b>	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
<b>CG</b>	Cromatografia Gasosa
<b>CLAE</b>	Cromatografia Líquida de Alta Eficiência
<b>CSV</b>	Valores Separados por Virgula
<b>EEM</b>	Matriz Excitação-Emissão de Fluorescência
<b>EMBRAPA</b>	Empresa Brasileira de Pesquisa Agropecuária
<b>EUA</b>	Estados Unidos da América
<b>FTIR</b>	Espectroscopia na Região do Infravermelho com Transformada de Fourier
<b>GA</b>	Algoritmo Genético
<b>HCA</b>	Análise de Agrupamentos Hierárquicos
<b>ICP-MS</b>	Espectrometria de Massa por Plasma Acoplado Indutivamente
<b>LAQIA</b>	Laboratório de Quimiometria e Instrumentação Analítica
<b>LDA</b>	Análise Discriminante Linear
<b>MAPA</b>	Ministério da Agricultura, Pecuária e Abastecimento
<b>PCA</b>	Análise de Componentes Principais
<b>pH</b>	Potencial Hidrogeniônico
<b>PLS</b>	Regressão por Mínimos Quadrados Parciais
<b>PLS-DA</b>	Análise Discriminante por Mínimos Quadrados Parciais
<b>PMT</b>	Tubo Fotomultiplicador
<b>PPGQ</b>	Programa de Pós-Graduação em Química
<b>PTR-MS</b>	Espectrometria de Massas de Reação de Transferência de Prótons
<b>QAV</b>	Química Analítica Verde
<b>RMSEV</b>	Raiz Quadrada do Erro Médio Quadrático de Validação
<b>RS</b>	Estado do Rio Grande do Sul
<b>SIMCA</b>	Modelagem Independente por Analogia de Classe
<b>SPA</b>	Algoritmo de Projeções Sucessivas
<b>SW</b>	Algoritmo Stepwise
<b>TCC</b>	Taxa de Classificação Correta

<b>UFRGS</b>	Universidade Federal do Rio Grande do Sul
<b>UNISC</b>	Universidade de Santa Cruz do Sul
<b>UNLPam</b>	Universidad Nacional de La Pampa
<b>UV-VIS</b>	Espectroscopia na Região do Ultravioleta-Visível

## RESUMO

O consumo de suco de uva vem crescendo dia a dia em razão da procura constante dos consumidores que buscam obter uma dieta mais saudável pela inclusão de alimentos naturais. O Estado do Rio Grande do Sul (RS) é responsável por mais de 50% do volume de colheita de uva e em torno de 90% da produção de sucos de uva no Brasil, o que ressalta a importância de estudos que garantam a autenticidade e procedência do alimento. O desenvolvimento de metodologias analíticas que visam auxiliar na detecção de possíveis adulterações e/ou falsificações intencionais do produto apresenta-se pertinente, especialmente aos sucos orgânicos que possuem maior valor comercial agregado. Neste contexto, a presente tese desenvolveu novas metodologias analíticas para a identificação e classificação de sucos de uva integral produzidos na região serrana do RS com relação ao sistema de produção agrícola implementado (convencional e orgânico), utilizando dados espectrais de técnicas espectroscópicas combinadas com ferramentas quimiométricas e seleção de variáveis. Os melhores resultados obtidos demonstraram que a análise discriminante linear (LDA) combinada com diferentes algoritmos de seleção de variáveis pode classificar com alta acurácia as amostras de suco de uva integral de acordo com o processo de produção com 100% no conjunto de treinamento do modelo e 100% na predição das classes das amostras do conjunto de teste. Além disso, em geral, os melhores modelos apresentaram elevada sensibilidade e seletividade para as amostras de treinamento e de teste. A seleção das variáveis mais significativas foi determinada para cada conjunto de dados e permitiu explicar a separação das classes. O uso de técnicas espectroscópicas associadas com procedimentos quimiométricos possibilitou o desenvolvimento de metodologias analíticas eficientes e confiáveis para a classificação de sucos de uva integral e que podem ser aplicadas para o controle de qualidade em indústrias alimentícias.

**Palavras-chave:** Suco de Uva Integral, Sistema de Produção Agrícola, Classificação, Espectroscopia, Quimiometria.



## ABSTRACT

The consumption of grape juice has been growing day by day due to the constant demand of consumers who seek to obtain a healthier diet through the inclusion of natural foods. The State of Rio Grande do Sul (RS) is responsible for more than 50% of the grape harvest volume and around 90% of the grape juice production in Brazil, which highlights the importance of studies that guarantee the authenticity and origin of the food. The development of analytical methodologies that aim to assist in the detection of possible adulterations and/or intentional falsifications of the product is relevant, especially for organic juices that have greater added commercial value. In this context, this thesis developed new analytical methodologies for the identification and classification of whole grape juice produced in the mountainous region of RS in relation to the agricultural production system implemented (conventional and organic), using spectral data from spectroscopic techniques combined with chemometric tools and selection of variables. The best results obtained showed that the linear discriminant analysis (LDA) combined with different variable selection algorithms can classify with high accuracy the grape juice samples according to the production process with 100% in the model training set and 100% in predicting classes of test set samples. Furthermore, in general, the best models showed high sensitivity and selectivity for training and test samples. The selection of the most significant variables was determined for each data set and made it possible to explain the separation of classes. The use of spectroscopic techniques associated with chemometric procedures enabled the development of efficient and reliable analytical methodologies for the classification of whole grape juices and that can be applied for quality control in food industries.

**Keywords:** Whole Grape Juice, Agricultural Production System, Classification, Spectroscopy, Chemometrics.

## 1. INTRODUÇÃO

O consumo de alimentos naturais e funcionais vem crescendo a cada ano e é considerado fator essencial para promoção de benefícios à saúde humana com a garantia de boa nutrição e a prevenção de doenças crônicas não transmissíveis, incluindo hipertensão arterial, cardiovasculares e câncer<sup>1,2</sup>. Uma dieta balanceada com a inserção de legumes, verduras, frutas e seus derivados é fundamental, uma vez que a conscientização por um estilo de vida mais saudável está subordinada a hábitos alimentares mais nutritivos<sup>3</sup>.

Nessa direção, os sucos de frutas são fontes ricas de vitaminas, carboidratos, carotenoides e minerais que facilitam o bom funcionamento do organismo, além de serem apreciados por aspectos sensoriais como sabor, frescor e aroma<sup>3</sup>. Em razão das vantagens nutricionais dessas bebidas, o fomento a pesquisas tem aumentado de forma significativa, o que desponta em favor da população que está mais atenta ao que consome<sup>4,5</sup>.

A uva *in natura* e seus produtos derivados como geleias, vinhos e sucos são alimentos bastante consumidos em todo o mundo, pois estão disponíveis comercialmente pela elevada produtividade que o setor da viticultura assegura e apresentam preço razoável para compra. Além disso, a fruta e seus derivados têm recebido grande interesse por possuírem propriedades funcionais intrínsecas relacionadas aos bioativos presentes, especialmente os compostos polifenólicos como antocianinas, ácidos fenólicos e taninos<sup>6,7</sup>.

Estudos apontam que o consumo frequente de suco de uva é benéfico ao bem-estar do ser humano, devido que estas substâncias apresentam qualidades cardioprotetoras, anti-inflamatórias, antioxidantes, entre outras<sup>8-13</sup>. Ainda, nesse contexto de alimentos ditos saudáveis, outro destaque tem sido o aumento na busca pelo consumo dos produtos orgânicos, que não utilizam agrotóxicos, procuram usar o mínimo de fertilizantes orgânicos na produção agrícola e tem práticas de gestão voltadas para mitigar impactos ambientais<sup>14-16</sup>.

O suco de uva integral pode ser obtido por vários métodos de extração, sem o acréscimo de açúcares e na sua concentração natural, dependendo do volume de processamento e da tecnologia disponível<sup>17-19</sup>. O controle de qualidade dos sucos de uva está estabelecido em legislação em que se avalia parâmetros físico-químicos, composição química e aspectos sensoriais, sobretudo utilizando análises clássicas e instrumentais<sup>20,21</sup>.

Entretanto, esses métodos padronizados, apesar de serem precisos, são geralmente caros, demorados e necessitam experiência de operação<sup>22,23</sup>.

O desenvolvimento de metodologias analíticas confiáveis, de rápida obtenção de resultados, práticos, e, de preferência, não destrutivos e de baixo custo é muito desejado para garantir a autenticidade dos alimentos<sup>22-24</sup>. Nesse quadro inserem-se o emprego de técnicas de espectroscopia molecular, intensamente exploradas nos últimos anos na análise de produtos alimentares<sup>25-27</sup>, pois fornecem valiosas informações sobre a interação da radiação eletromagnética com os diferentes compostos presentes na amostra.

O aperfeiçoamento das tecnologias na instrumentação analítica na Química e o desenvolvimento dos microcomputadores têm permitido a produção e o processamento de informações em maior volume e rapidez. No entanto, a interpretação e análise não é tarefa simples e requer a utilização de ferramentas para tratamento desse grande conjunto de dados. Dentro desse cenário se aplica a Quimiometria, que consiste no uso de princípios estatísticos, modelos matemáticos e ciência da computação para manusear, analisar e prever dados de origem química. Em outras palavras, permite converter dados numéricos adquiridos por meio dos instrumentos em informação de caráter químico<sup>28,29</sup>. Portanto, as técnicas de análise multivariada são procedimentos úteis e aplicadas com sucesso na ciência dos alimentos em diversas abordagens<sup>30-32</sup>.

## 1.1 OBJETIVO GERAL

O objetivo geral dessa tese foi desenvolver novas metodologias analíticas para a classificação de sucos de uva integral com respeito ao sistema de produção agrícola convencional e orgânico, utilizando os dados espectrais de técnicas espectroscópicas aliada às ferramentas quimiométricas, para construção de modelos por análise discriminante linear (LDA) e análise discriminante por mínimos quadrados parciais (PLS-DA).

### 1.1.1 Objetivos Específicos

- Examinar os conjuntos de dados obtidos pelas técnicas de espectroscopia na região do visível, espectroscopia na região do infravermelho médio e espectroscopia de fluorescência em amostras de suco de uva a fim de analisar o poder discriminante dos mesmos;

- Avaliar a capacidade preditiva das técnicas supervisionadas de classificação multivariada empregadas LDA e PLS-DA;
- Comprovar a eficiência e a confiabilidade dos métodos analíticos desenvolvidos por meio das figuras de mérito de sensibilidade, especificidade e acurácia;
- Selecionar as informações mais relevantes dos conjuntos de dados por meio de diferentes algoritmos de seleção de variáveis a fim de estabelecer distinção entre as duas classes;
- Classificar as amostras de sucos de uva conforme o sistema agrícola de produção (convencional e orgânico).

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 BREVE HISTÓRICO DA VITICULTURA MUNDIAL E BRASILEIRA

O cultivo da uva para fins de consumo é tão remoto quanto o início da civilização humana. Vestígios mais antigos de fósseis de folhas e de sementes de uva são datados de 7000 a 5000 anos a.C. e foram descobertos onde atualmente corresponde ao país da Geórgia, localizado na região do Cáucaso, situado entre a Europa oriental e Ásia ocidental. Territórios adjacentes como dos antigos povos da Mesopotâmia e do Egito remontam a existência de viticultura há 5000 anos por meio de registros arqueológicos<sup>33-35</sup>. Os fenícios também contribuíram com a expansão da uva por meio do comércio marítimo na região da Grécia, vindo esta a incorporá-la rapidamente em sua cultura<sup>34</sup>.

O Império Romano foi um dos grandes responsáveis pela disseminação da uva pela Europa, primeiro nas zonas costeiras do Mediterrâneo e posteriormente em direção ao interior, acompanhando as principais rotas de comércio da época<sup>33-35</sup>. Relatos históricos revelam que gregos e romanos concentravam o mosto da uva por aquecimento para conservação e posterior uso na produção de sucos e doces, para adoçar vinhos e ser um substituto do mel<sup>36</sup>.

O início do cultivo da uva no Brasil foi realizado por Martim Afonso de Sousa em 1532, o qual trouxe as primeiras videiras de *Vitis vinífera* provenientes de Portugal e da Espanha. O local de plantio foi a capitania hereditária de São Vicente, atual Estado de São Paulo, em que se pressupõe que tinham a finalidade para produção de vinho. No começo, a cultura foi cultivada no litoral, no entanto se dispersaram pelas demais regiões devido a boa adaptação ao solo e ao clima fornecendo bons frutos. As primeiras videiras plantadas no Estado do Rio Grande do Sul (RS) datam de 1626 pelos padres jesuítas e, mais tarde, foram trazidas também pelos imigrantes alemães. Entretanto, em 1789, um decreto protecionista econômico promulgado por Portugal vetou o plantio de uva para produção e comercialização de vinhos no país<sup>33-36</sup>.

Durante o século XIX decorreu o declínio da plantação de uva em razão da importação de videiras infectadas com doenças fúngicas. Mantida como agricultura familiar, a viticultura recebeu um forte estímulo a partir de 1875 com a vinda dos imigrantes italianos ao Brasil, especialmente no RS, que trouxeram consigo espécies europeias (*Vitis vinífera*) e a tradição do cultivo da uva para a produção e consumo de vinhos. Contudo, as condições climáticas do Rio Grande do Sul não favoreciam um

plantio favorável dessas castas. A inserção de espécies de origem americana (*Vitis labrusca*) forneceu a base para o estabelecimento e consolidação do cultivo de uva no Brasil, especialmente no Rio Grande do Sul, para fins de comércio. A viticultura no Brasil é recente em comparação com países tradicionais como Itália, França, Espanha e Portugal<sup>33-35,37</sup>.

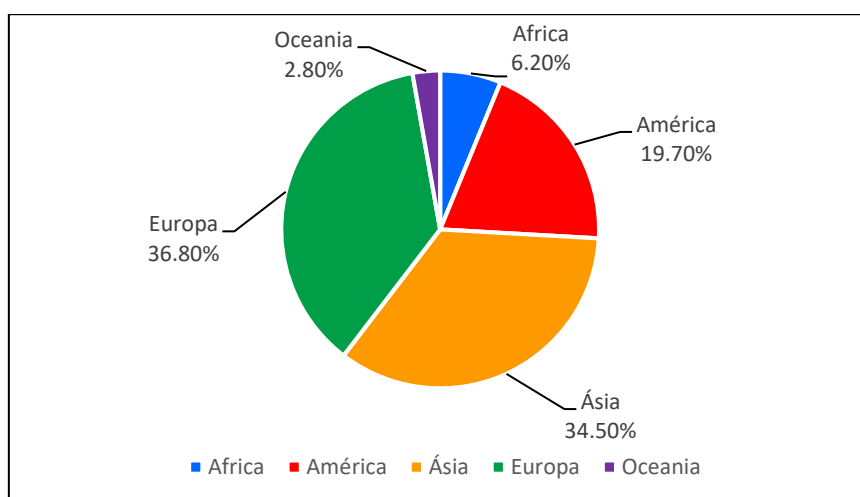
A viticultura brasileira exhibe particularidades e complexidades que a diferencia das tradicionais regiões da cultura no mundo, relacionadas a condições territoriais, ambientais e históricas. As áreas agrícolas de cultivo da uva no Brasil ficam localizadas principalmente nas regiões Sul, Sudeste e Nordeste e essas se distinguem pelo solo, condições climáticas, modo de plantio, processamento, ciência e engenharia aplicadas, entre outros fatores. Esses fatores de diversidade conduzem a ciclos típicos das temporadas de colheita devido aos estágios vegetativos das videiras e, como consequência, originam produtos com propriedades e características próprias<sup>38</sup>.

## 2.2 SITUAÇÃO ATUAL DA PRODUÇÃO DE UVAS E SUCOS DE UVA

Os maiores países produtores de uva atualmente são a China, a Itália, a Espanha, a França, os Estados Unidos, a Turquia, a Índia e o Chile. Na última década, surgiram como potenciais produtores de uva a Argentina, a África do Sul, o Irã, o Uzbequistão, a Austrália e o Brasil<sup>24,39,40</sup>. A Figura 1 ilustra a produção média de uva por continente no período de 2010 a 2020, em que a Europa (36,8%) e a Ásia (34,5%) são as principais regiões. No mundo, a área plantada vem diminuindo ao longo das últimas décadas, totalizando em 2020 cerca de 7 milhões de hectares<sup>39</sup>.

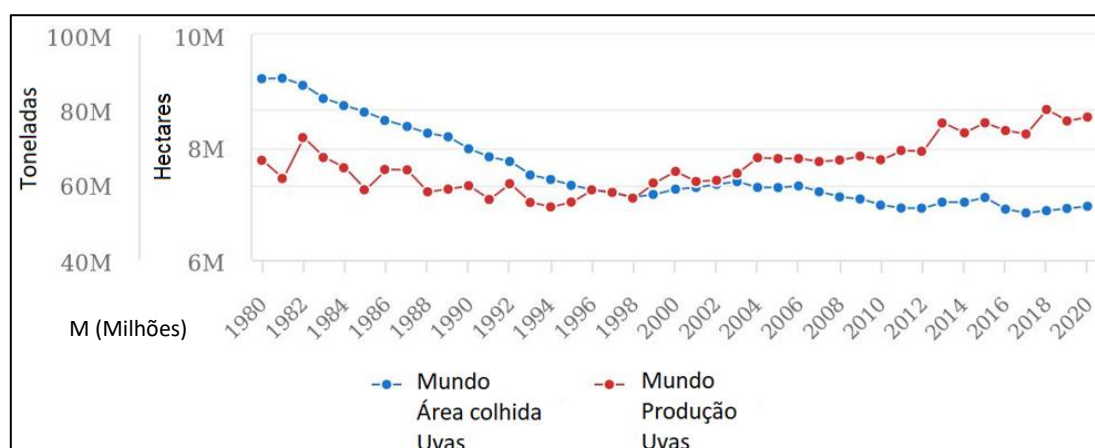
Entretanto, a produção de uvas segue fluxo inverso, com crescimento constante desde o final dos anos 90 até os dias atuais, apresentando em 2020 um total de 78 milhões de toneladas produzidas<sup>39</sup>. Esse fenômeno muito se deve aos avanços nas pesquisas em ciências e engenharias de alimentos na busca por melhora na produtividade. A Figura 2 retrata a evolução da quantidade produzida de uva em função da área colhida no mundo entre os anos de 1980 a 2020. Sobre a produção de suco de uva ao redor do mundo, é estimado, atualmente, em torno de 11 a 12 milhões de hectolitros<sup>24,40</sup>.

**Figura 1.** Produção média de uva por continente no período de 2010 a 2020.



Fonte: adaptada de FAO, 2022<sup>39</sup>.

**Figura 2.** Evolução da quantidade produzida e área de colheita de uva no mundo no período de 1980 a 2020.

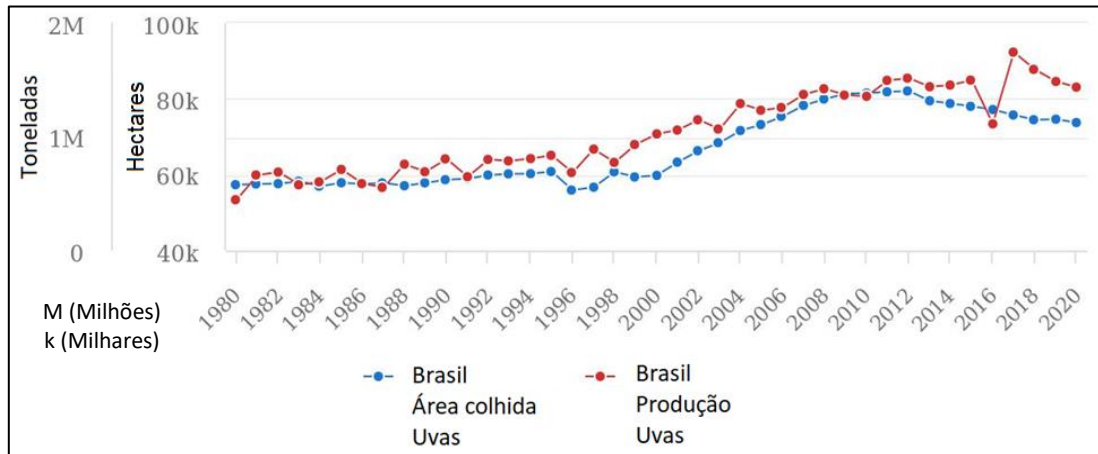


Fonte: adaptada de FAO, 2022<sup>39</sup>.

No Brasil, a produção média de uvas permanece ao redor de 1,5 milhões de toneladas por ano e é destinada ao mercado doméstico e internacional<sup>39,41</sup>. Deste volume, a maior parte produzida é reservada ao processamento via artesanal ou industrial para a elaboração de vinhos, sucos, geleias e outros derivados, enquanto que o restante da produção é comercializado para o consumo *in natura* como uva de mesa<sup>42,43</sup>. Os produtos industrializados são elaborados a partir de uvas de origem americana, especialmente cultivares de *Vitis labrusca*, *Vitis bourquina* e diversos híbridos interespecíficos<sup>44</sup>. Em 2020 o consumo médio por habitante de uva de mesa (consumo *in natura* e doces) foi de 3,49 kg<sup>42</sup>. A Figura 3 retrata a evolução da quantidade produzida de uva em função da área colhida no Brasil entre os anos de 1980 a 2020, em que o ano de 2016 apresentou uma

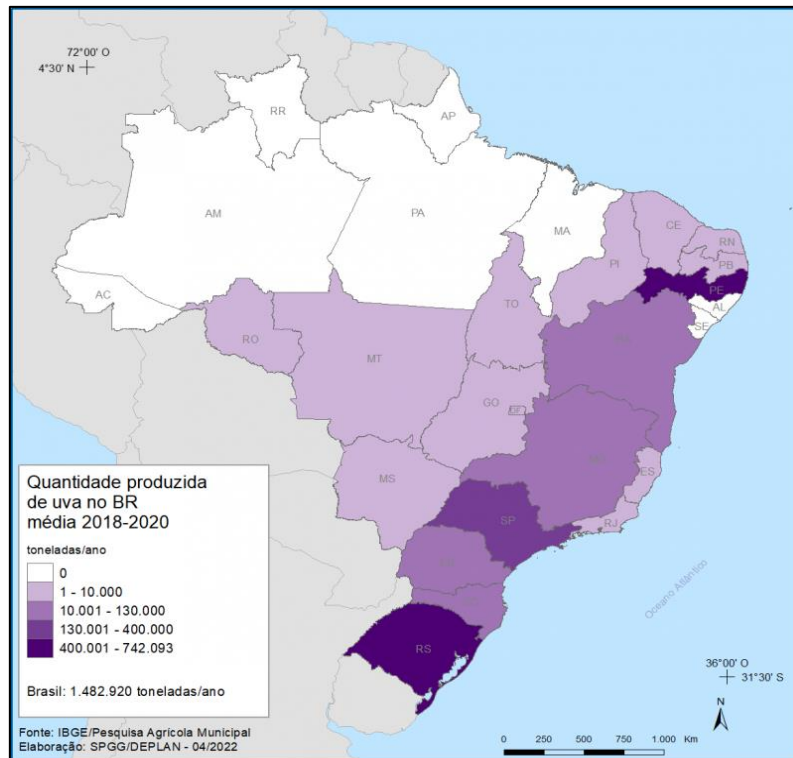
queda na produção de uvas devido a fatores climáticos adversos (geadas, chuvas excessivas e granizo) que comprometeram o desenvolvimento das videiras<sup>45</sup>. A Figura 4 mostra a quantidade média produzida de uva no período de 2018 a 2020, em que os Estados com coloração mais escura são os que mais elaboraram.

**Figura 3.** Evolução da quantidade produzida e área de colheita de uva no Brasil no período de 1980 a 2020.



Fonte: adaptada de FAO, 2022<sup>39</sup>.

**Figura 4.** Quantidade média produzida de uva pelos estados brasileiros no período de 2018 a 2020.



Fonte: Atlas Socioeconômico do RS, 2022<sup>43</sup>.



O cultivo de uva está concentrado principalmente nas Regiões Sul e Nordeste do país<sup>43</sup>. A Região Sul é responsável por mais de 60% da produção de uva no país, com destaque para o RS que produz cerca de 50%, sendo o maior produtor nacional da fruta, e fomenta cerca de 90% da fabricação nacional de vinhos e sucos de uva. Na última década, a região Nordeste, especialmente a região do Vale do São Francisco, vem apresentando uma produção significativa voltada à elaboração de suco e uva. Em 2022, a área de colheita de uva apresentou um total de 75 mil hectares plantados, a maioria instalada em pequenas propriedades de agricultura familiar<sup>38,42</sup>. A Tabela 1 mostra o histórico de anos recentes da área plantada e da quantidade produzida de uva no Brasil e no Rio Grande do Sul, em que no País houve diminuição da área à medida que o volume da produção se manteve estável, enquanto que no RS a correlação se manteve similar.

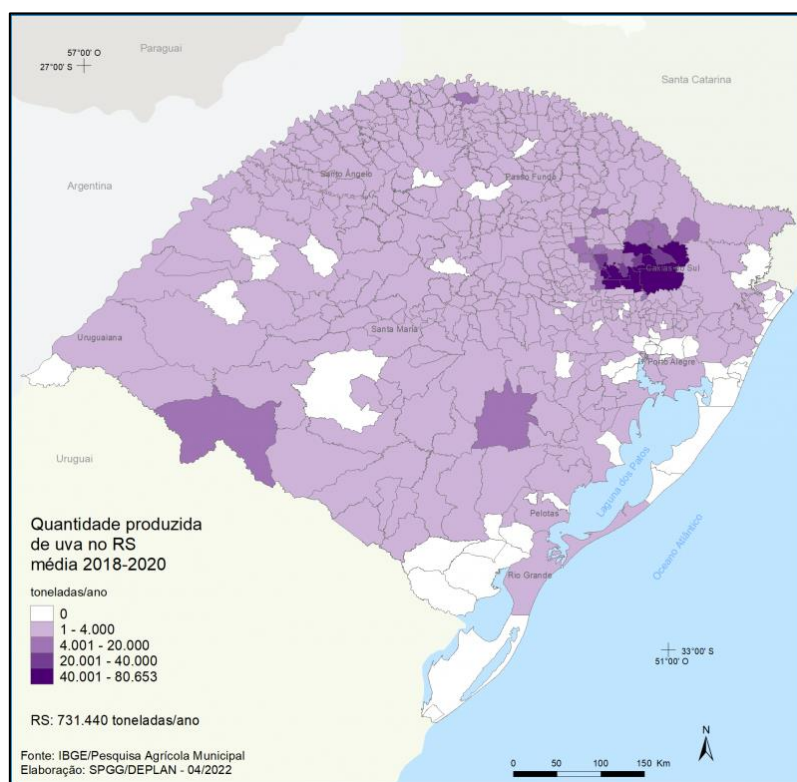
**Tabela 1.** Área plantada e quantidade produzida de uva no Brasil e no Rio Grande do Sul no período de 2008 a 2022.

Ano	Brasil		Rio Grande do Sul	
	Área Plantada (Hectares)	Quantidade Produzida (Toneladas)	Área Plantada (Hectares)	Quantidade Produzida (Toneladas)
2008	81.286	1.421.431	47.206	776.964
2009	81.677	1.365.491	48.259	737.363
2010	81.534	1.355.461	48.753	694.518
2011	81.840	1.495.336	49.198	839.286
2012	82.897	1.514.768	50.180	840.251
2013	79.759	1.439.535	50.056	807.693
2014	78.779	1.454.183	50.007	812.517
2015	78.026	1.497.302	49.739	876.215
2016	77.299	1.113.345	49.226	413.735
2017	75.968	1.743.430	48.397	965.887
2018	74.480	1.592.806	46.858	823.698
2019	74.852	1.485.806	46.671	667.239
2020	74.435	1.435.596	46.174	735.342
2021	75.086	1.702.660	46.809	951.258
2022	75.595	1.501.555	47.034	734.994

Fonte: adaptada de IBGE, 2022<sup>41</sup>.

A origem da cultura permanente de uva no RS é decorrência da forte influência da colonização italiana em meados do século XIX e está concentrada sobretudo na região serrana localizado no nordeste do Estado. Os municípios de Bento Gonçalves e de Flores da Cunha se destacam por apresentarem uma produção média anual acima de 80 mil toneladas por ano. Mais recentemente, as regiões localizadas a sudoeste do Estado como a Fronteira Oeste, a Campanha e o Médio Alto Uruguai vêm se desenvolvendo na elaboração de uva para a viticultura<sup>43</sup>. A Figura 5 mostra a quantidade média produzida de uva no RS no período entre 2018 e 2020.

**Figura 5.** Quantidade média produzida de uva pelos municípios do RS no período de 2018 a 2020.



Fonte: Atlas Socioeconômico do RS, 2022<sup>43</sup>.

Em relação à produção de suco de uva integral pelo sistema agrícola convencional, o Rio Grande do Sul na safra 2021 elaborou um volume maior de 68 milhões de litros, enquanto que o suco produzido pelo sistema agrícola orgânico correspondeu a mais de 700 mil litros, conforme demonstrado na Tabela 2<sup>46</sup>. Já a comercialização de suco de uva produzido no RS esteve ao redor de 145 milhões de litros registrado para o ano de 2020, correspondendo a um consumo per capita de 1,36 litros<sup>42</sup>.

**Tabela 2.** Evolução da produção de sucos de uva nas safras 2017 a 2021 no RS.

Suco de Uva	Quantidade produzida (Litros)				
	2017	2018	2019	2020	2021
Adoçado	177.774	152.446	120.320	107.289	40.450
Concentrado	30.059.271	31.099.910	26.861.009	27.963.865	32.131.219
Integral	46.839.961	34.569.601	50.569.601	40.718.523	68.038.479
Total origem convencional	77.136.206	65.821.957	77.221.096	68.789.677	100.210.147
Total origem orgânico	Não disponível	Não disponível	628.474	469.910	722.117

Fonte: adaptada de RS, 2021<sup>46</sup>.

### 2.3 SISTEMAS CONVENCIONAL E ORGÂNICO DE PLANTIO DE UVA

Em muitos países, como no Brasil, o cultivo de uvas está estabelecido por dois processos de produção: o convencional e o orgânico. O sistema agrícola convencional é o manejo mais difundido no mundo, inclusive no Brasil a partir da década de 1970, em decorrência da “Revolução Verde” por meio de inovações tecnológicas<sup>47</sup>. O método é aplicado principalmente nas plantações latifundiárias e é caracterizado pela aplicação intensiva de agrotóxicos (pesticidas e herbicidas sintéticos), fertilizantes e adubos químicos, vasto uso de mecanização e maquinário, produção em larga escala, dirigido à monocultura e utilização de plantas modificadas geneticamente (transgênicos)<sup>47-50</sup>.

O processo emprega substâncias químicas potencialmente poluentes com os objetivos de eliminar vários tipos de pragas danosas às lavouras e elevar a produtividade com menor custo de operação e maximização de lucros. No entanto, quando os insumos são usados de modo inadequado, a agricultura convencional pode ocasionar consideráveis impactos e prejuízos ao meio ambiente e à saúde humana, tanto entre os trabalhadores rurais expostos diretamente aos produtos químicos, quanto aos consumidores que acabam ingerindo essas substâncias tóxicas<sup>47-50</sup>.

O modelo de agricultura orgânico segue algumas diretrizes, como a Lei Federal nº 10.831/2003, que traz o regramento para a produção e comercialização dos produtos orgânicos no Brasil, e o Decreto nº 6.323/2007, que confere a regulamentação. Segundo a legislação, “considera-se sistema orgânico de produção agropecuária todo aquele em que se adotam técnicas específicas, mediante a otimização do uso dos recursos naturais e

socioeconômicos disponíveis e o respeito à integridade cultural das comunidades rurais, tendo por objetivo a sustentabilidade econômica e ecológica, a maximização dos benefícios sociais, a minimização da dependência de energia não-renovável, empregando, sempre que possível, métodos culturais, biológicos e mecânicos, em contraposição ao uso de materiais sintéticos, a eliminação do uso de organismos geneticamente modificados e radiações ionizantes, em qualquer fase do processo de produção, processamento, armazenamento, distribuição e comercialização, e a proteção do meio ambiente”<sup>51</sup>.

Portanto, o processo de produção orgânico é identificado pela não utilização de agrotóxicos, aditivos químicos ou outros produtos fitossanitários, dirigido a rotação de culturas, manejo do solo usando matéria orgânica vegetal ou animal (húmus de minhoca, esterco curtido, compostagem, entre outros), aplicação de minerais naturais, controle biológico de pragas e não uso de transgênicos<sup>52,53</sup>.

O sistema de cultivo é considerado sustentável, por adotar práticas de gestão que respeitam os ecossistemas amenizando impactos ambientais. A produção em larga escala e as constantes adulterações ou falsificações intencionais pela alteração parcial ou total da matéria-prima original por outra de menor valor agregado são desafios e dificuldades encontrados no sistema de cultivo orgânico, que tem como consequência o encarecimento do produto final aos consumidores<sup>54</sup>. Apesar dessas limitações, é considerável a expansão por pequenos produtores rurais nesse processo de plantio a fim de atender à gradual demanda de mercado<sup>55,56</sup>.

Na produção orgânica de uvas a adversidade é grande pois a videira, se deixada sob condições naturais, é suscetível a diversas doenças que podem levar a perda de frutos. Diversos fatores devem ser levados em consideração para a produção de uva orgânica como a escolha da área (solo, orientação solar, declive, ventilação), preparo da área (adubação de correção dos nutrientes), adubações de crescimento e manutenção durante o crescimento da planta, implantação de plantas de cobertura (reciclagem dos nutrientes) e controle de pragas (agentes biológicos como caldas, óleos vegetais e extratos de plantas)<sup>55</sup>.

## 2.4 CARACTERÍSTICAS DA UVA E DAS VIDEIRAS

A uva é um fruto de origem asiática do tipo baga, de tamanho, formato, aroma e consistência variáveis, apresentando pigmentação verde, vermelho, preto ou roxo. A polpa é comestível e contém até quatro sementes, apresentando sabor doce, ácido, amargo ou

adstringente. Os principais produtos derivados da uva são o vinho, suco, licor, geleia e demais doces, além de fornecer subprodutos, como corantes naturais, ácido tartárico, óleo de semente e taninos. É uma fruta típica de regiões com clima temperado e do tipo não-climatérico, ou seja, não amadurece após a colheita, devendo ser colhida no ponto ideal de maturação<sup>20,57,58</sup>.

A videira (*Vitis sp.*), da família *Vitaceae*, é uma planta trepadeira lenhosa, com gavinhas de fixação, que exibem troncos tortuosos e retorcidos, podendo atingir até sete metros de comprimento. O caule é espesso e resistente, com coloração verde quando jovem, tornando-se escuro posteriormente. As folhas são grandes, verdes, palmadas e com cinco lobos. As videiras compreendem inúmeras espécies e seus híbridos (cruzamentos). Comercialmente, são classificadas basicamente em dois grupos: as finas (cultivar europeia de *Vitis vinífera L.*) e as rústicas (cultivares americanas – *Vitis labrusca* e *Vitis bourguina* – e híbridos interespecíficos – *Vitis spp.*)<sup>20,57,58</sup>.

As videiras reservadas à elaboração de suco de uva devem possuir alguns atributos como atingir uma boa relação açúcar/acidez, rendimento em mosto, ter agradável aroma e sabor e apropriado nível de amadurecimento do fruto<sup>20,57,58</sup>. A Figura 6 ilustra modelo típico do plantio de videira em pérgola ou latada, técnica de condução mais utilizada no RS que consiste em guiar as plantas em estruturas elevadas. Esse sistema permite que as plantas recebam uma maior exposição à luz solar e melhor ventilação, o que pode resultar em maior produção de uvas e melhor qualidade da safra<sup>59</sup>.

**Figura 6.** Modelo típico do cultivo de videira em pérgola ou latada.



Fonte: Embrapa, 2017<sup>60</sup>.

## 2.5 DEFINIÇÃO E CLASSIFICAÇÃO DO SUCO DE UVA

De acordo com o Decreto nº 6.871/2009, suco ou sumo de fruta é a bebida líquida límpida ou turva de cor, aroma e sabor característicos, não fermentada, não alcoólica, não concentrada e nem diluída extraída da fruta fresca, madura e sã, ou parte do vegetal de origem, por meio de processo tecnológico adequado, submetido a tratamento que assegure a sua apresentação e conservação até o momento do consumo. O suco não poderá conter substâncias estranhas à fruta, salvo previsão legal. Além disso, é proibida a adição de aromas e corantes artificiais<sup>61</sup>. A Lei nº 7.678/1998, dispõe que “o suco de uva é a bebida não fermentada, obtida do mosto simples, sulfitado ou concentrado, de uva sã, fresca e madura”<sup>62</sup>.

O suco de uva é um alimento complexo e variado de substâncias químicas, sendo rico em metabólitos primários como proteínas, carboidratos e lipídios, e metabólitos secundários, com destaque para os compostos polifenólicos como antocianinas, taninos, flavanóis, flavonóis, catequinas e outros flavonoides. Estes compostos contribuem para as propriedades sensoriais, particularmente em termos de cor, amargura e adstringência. Em geral, o teor de compostos fenólicos aumenta ao longo do amadurecimento da uva e pode ser diretamente alterado pelo déficit hídrico<sup>22,63</sup>.

O suco de uva é destacado por ser uma bebida única, tanto segundo o aspecto nutritivo e energético quanto pelos aspectos funcional e terapêutico. É um alimento de paladar doce e ácido, simultaneamente, apresentando fácil digestão e absorção dos seus componentes pelo organismo humano. Ainda, é considerada uma bebida não fermentada e não alcoólica, no qual o teor de álcool deve ser sempre inferior a 0,5% v/v<sup>20,35,57,64</sup>.

O suco de uva pode ser classificado como tinto, branco e rosé ou rosado com relação a variedade utilizada, o método de elaboração aplicado e ao aspecto visual final. Por ser um produto natural, possui características intrínsecas de sabor e aroma que são inerentes ao varietal de uva empregado. Segundo a legislação brasileira, o suco de uva para consumo é classificado como adoçado, concentrado, desidratado, gaseificado, integral e reconstituído. Essa categorização se deve pelo tipo de processo utilizado e pelas suas características, conforme mostrado na Tabela 3<sup>20,35,57,64</sup>.

**Tabela 3.** Classificação dos sucos de uva processados segundo a legislação brasileira.

Tipo de suco de uva	Características
Adoçado	É permitida a adição de açúcar, como a sacarose, na quantidade máxima de um décimo em peso dos açúcares do mosto.
Concentrado	Parcialmente desidratado para elevar em, no mínimo, 50% o teor dos sólidos solúveis presentes no respectivo suco integral. Proibido a adição de açúcares.
Desidratado	Obtido a partir do suco integral. Apresenta apenas 3% de umidade e encontra-se na fase sólida. Proibido a adição de açúcares.
Gaseificado	Adicionado de dióxido de carbono, de 1,1 até 3 atmosferas, a 20 °C.
Integral	Apresentado na sua concentração e composição natural, límpido ou turvo, sem adição de açúcares, corantes ou aromas, apenas os que provém da uva.
Reconstituído	Obtido pela diluição do suco concentrado ou desidratado até a concentração original do suco integral ou ao teor mínimo de sólidos solúveis do suco integral.

Fonte: Brasil, 2018<sup>64</sup>.

A elaboração de suco é uma das diversas possibilidades de beneficiamento da uva, onde no Brasil corresponde em torno de 10% da produção total da fruta. Em virtude da simplicidade de elaboração, com qualquer variedade de uva sendo capaz de ser empregada desde que demonstre estar em boa condição sanitária e na maturação apropriada, aliada às propriedades organolépticas e importância nutricional, o suco de uva é considerado um alimento valioso e saudável que pode auxiliar na dieta alimentar<sup>20,35,57</sup>.

## 2.6 COMPOSIÇÃO QUÍMICA DO SUCO DE UVA

A composição química e valor nutricional do suco de uva são similares ao da própria uva, com exceção da quantidade maior de fibra bruta e óleo vistos nas sementes. Na sua constituição estão todos os compostos essenciais, sendo rico em açúcares, sais

minerais, ácidos orgânicos, vitaminas e compostos fenólicos, e, ao mesmo tempo, apresentando baixo teor em proteínas e lipídios<sup>20,57</sup>. A composição química está atrelada a inúmeros fatores como a uva empregada (varietal), ao grau de maturação das bagas, à constituição do solo, ao estresse hídrico, à exposição solar do cacho, ao sistema de cultivo, aos procedimentos agrônômicos, patogênese, método de prensagem, aos tratamentos a que a uva foi submetida, entre outros<sup>6,65,66</sup>. As modificações físicas e bioquímicas que ocorrem durante o ciclo de desenvolvimento da fruta dependem do estágio de amadurecimento da uva, sendo a maturação um dos principais aspectos relacionados a qualidade do suco, em conjunto com o estado sanitário da baga<sup>20,57</sup>. A seguir, a Tabela 4 apresenta os valores nutricionais para o suco de uva e a Figura 7 estão representados alguns dos componentes químicos, na sua forma estrutural, presentes no suco de uva.

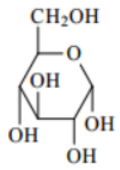
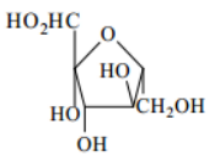
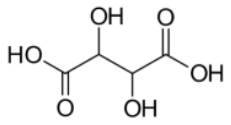
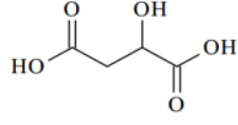
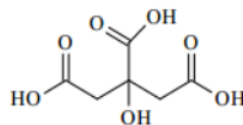
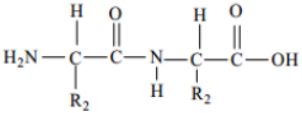
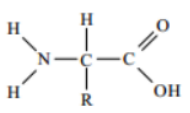
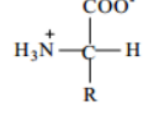
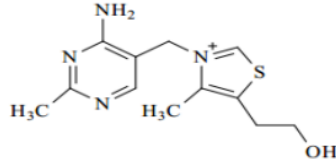
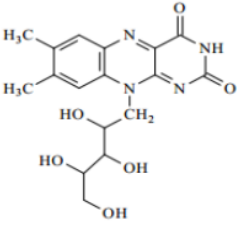
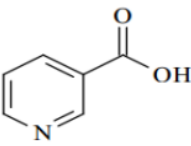
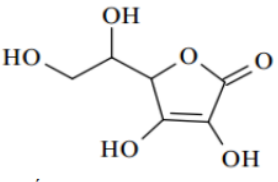
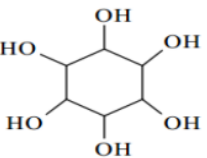
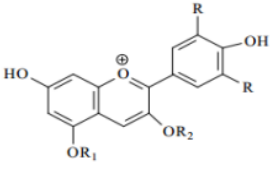
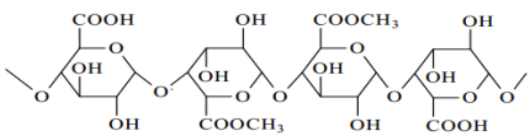
**Tabela 4.** Valores nutricionais do suco de uva.

Princípio alimentar	Concentração	Princípio alimentar	Concentração
Água (%)	81 – 86	Aminoácidos (g/L)	0,6 – 2,0
Calorias (cal/L)	700 – 900	Inositol (mg/100 g)	40 – 50
Açúcares (g/L)	140 – 180	Tiamina (µg/100 g)	50 – 60
Minerais (g/L)	1,5 – 3,0	Riboflavina (µg/100 g)	50 – 60
Lipídios (g/L)	1,0 – 2,0	Niacina (µg/100 g)	0,4 – 0,6
Pectina (g/L)	0,3 – 0,6	Ácido ascórbico (mg/100 g)	0,2 – 4,0
Proteínas (g/L)	2,0 – 3,0	Riboflavina (µg/100 g)	50 – 60

Fonte: adaptada de Rizzon, 1998<sup>57</sup>.



**Figura 7.** Fórmula estrutural de alguns componentes químicos presentes no suco de uva.

Açúcares		
		
Glicose (C <sub>6</sub> H <sub>12</sub> O <sub>6</sub> )		Frutose (C <sub>6</sub> H <sub>12</sub> O <sub>6</sub> )
Ácidos orgânicos		
		
Ácido tartárico (C <sub>4</sub> H <sub>6</sub> O <sub>6</sub> )	Ácido málico (C <sub>4</sub> H <sub>6</sub> O <sub>5</sub> )	Ácido cítrico (C <sub>6</sub> H <sub>8</sub> O <sub>7</sub> )
Substâncias nitrogenadas		
		
Polipeptídeos	Aminoácidos	Proteína
Vitaminas		
		
Tiamina (C <sub>12</sub> H <sub>17</sub> N <sub>4</sub> OS <sup>+</sup> )		Riboflavina (C <sub>17</sub> H <sub>20</sub> N <sub>4</sub> O <sub>6</sub> )
		
Niacina (C <sub>6</sub> H <sub>5</sub> NO <sub>2</sub> )	Ácido ascórbico (C <sub>6</sub> H <sub>8</sub> O <sub>6</sub> )	Inositol (C <sub>6</sub> H <sub>12</sub> O <sub>6</sub> )
Compostos fenólicos	Outros	
		
Antocianinas	Pectina	

Fonte: Venquiaruto, 2018<sup>36</sup>.

Os principais constituintes encontrados nos sucos de uva são os seguintes<sup>20,57</sup>:

**Água:** principal componente sob o aspecto quantitativo, é retirada do solo pelas raízes da videira e depositadas nas células da uva, passando para o suco durante o preparo.

**Açúcares:** representam os constituintes energéticos do suco, sendo a glicose e a frutose predominantes em proporções semelhantes. São facilmente absorvíveis pelo corpo humano. Estão relacionados com a espécie e o estado de maturação da uva.

**Ácidos orgânicos:** responsáveis pelo gosto ácido, estímulo da secreção salivar e do suco gástrico e pela capacidade bactericida do suco de uva. Os principais presentes são os ácidos tartárico, málico e cítrico.

**Minerais:** atuam na constituição dos ossos, sangue e nervos, além de assegurarem a alcalinidade do sangue e neutralizarem alguns ácidos, como o ácido úrico. Os principais encontrados são o potássio, sódio, cálcio, magnésio, manganês, ferro, fosfatos, sulfatos e cloretos. São absorvidos pela raiz da videira por meio da solução do solo e se acumulam nos frutos.

**Substâncias nitrogenadas:** são representados por proteínas, polipeptídeos, aminoácidos e nitrogênio amoniacal. Nos sucos de uva estão presentes os 20 aminoácidos considerados essenciais para o organismo humano.

**Compostos fenólicos:** as principais classes encontradas são os ácidos fenólicos, antocianinas e taninos. São responsáveis pela cor, estrutura e adstringência, além de apresentarem efeito antibiótico e ação benéfica de regulação da permeabilidade e da resistência dos vasos sanguíneos.

**Vitaminas:** são indispensáveis para o desenvolvimento e funcionamento de processos vitais do organismo como o metabolismo dos açúcares, conservação de resistência física e controle dos radicais livres, onde sua ausência ocasiona distúrbios e lesões. As principais encontradas são as vitaminas do complexo B (inositol, niacina, riboflavina e tiamina) e do complexo C (ácido ascórbico).

**Pectina:** auxilia no aumento da viscosidade devido ao seu poder gelificante e possui maior teor em relação aos demais sucos. É constituída por moléculas de ácido galacturônico, que podem ser hidrolisadas pelo calor ou pela atividade enzimática.

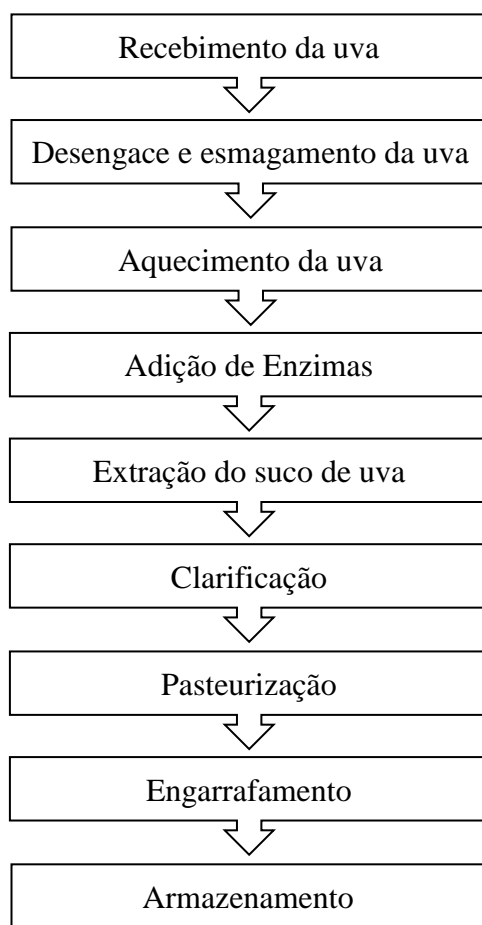
## 2.7 PROCESSO DE ELABORAÇÃO DO SUCO DE UVA

Os estudos de Louis Pasteur sobre processos de esterilização e de conservação em alimentos possibilitaram a expansão na produção e no consumo do suco de uva integral. O primeiro suco de uva processado nos Estados Unidos é datado no ano de 1869, quando o dentista Dr. Thomas Welch, auxiliado por sua esposa e filho, cozinharam uvas da variedade Concord, extraíram o seu suco e a engarrafaram imediatamente com rolhas de

cortiça. Para conservar o suco, aplicaram a teoria de Pasteur, em que as garrafas foram fervidas em água para evitar a fermentação e eliminar as leveduras presentes. A realização viabilizou o caminho para a industrialização do envase do suco de uva na América<sup>67</sup>.

Atualmente, o processo tecnológico aplicado na produção de suco de uva influencia na constituição química e nas propriedades do produto final, em particular o tempo e a temperatura de extração do suco, pois esses parâmetros regulam a intensidade e a solubilidade de difusão das substâncias presentes na casca para o mosto<sup>20,57</sup>. O suco de uva natural industrializado deve seguir um padrão de qualidade durante todo o processo de fabricação, que vai desde o recebimento da uva até ao engarrafamento. Em geral, o processo completo, desde o recebimento da fruta até o envase do suco, pode levar de alguns dias a algumas semanas, dependendo do tamanho da produção e dos equipamentos utilizados. A Figura 8 mostra as etapas envolvidas na elaboração de sucos de uva.

**Figura 8.** Etapas do processamento industrial para elaboração do suco de uva.



Fonte: Rizzon, 2007<sup>20</sup>.

O recebimento é a etapa em que a uva recém colhida chega na agroindústria e é realizado o controle do estado sanitário e peso, além da determinação do teor de açúcar do mosto para cálculo do rendimento. A operação de separação da ráquis é realizado numa desengaçadeira e é importante pois ela pode interferir negativamente com gosto amargo na composição do mosto. O esmagamento da uva tem a finalidade de extrair a cor, contribuindo na composição e aspecto do suco<sup>20,57</sup>.

A etapa de aquecimento da uva esmagada (65 a 90 °C) em termo-macerador tubular tem o objetivo de retirar da película algumas substâncias presentes como os compostos fenólicos, responsáveis pela cor. A adição de enzimas (catalisadores biológicos) comerciais auxilia no favorecimento de reações químicas, não alterando sabor e aroma, porém contribuem com maior intensidade na cor. As enzimas pectinases ou pectinolíticas são as mais usadas e ficam em contato com a uva esmagada por cerca de 1 a 2 horas à 60 °C para hidrolisar as pectinas, polissacarídeos estruturais presentes nos tecidos vegetais, favorecendo a extração e clarificação do suco de uva<sup>20,57</sup>.

A etapa da extração do suco de uva é realizada por meio dos equipamentos esgotador dinâmico e prensa descontínua com o intuito de separá-lo da fase sólida (sementes e casca). O suco obtido nesse processo tem o aspecto turvo pela presença de partes da película e sais de potássio em suspensão. A clarificação pode ser atingida por diferentes tratamentos como despectinização, filtração ou estabilização tartárica, a fim de deixar o suco de uva com aparência mais límpida e com menos precipitados<sup>20,57</sup>.

A conservação baseia-se na diminuição do contato do produto com o oxigênio atmosférico e na inibição do crescimento de micro-organismos, principalmente leveduras. Esse processo assegura a integridade do suco de uva até o consumo. A pasteurização, por meio de rápido aquecimento por tempo breve e imediato resfriamento a temperatura ambiente, é o método mais utilizado para esterilização do suco de uva. Análises físico-químicas de parâmetros analíticos para certificação conforme a legislação e avaliação sensorial de possíveis defeitos organolépticos são realizados no suco de uva antes do envase, geralmente em garrafas de vidro. Por fim, o armazenamento deve ser em local fresco, seco, arejado e protegido da luz solar direta<sup>20,57</sup>.

Existem várias metodologias para a produção de sucos de uva. O mais difundido entre as empresas produtoras no RS para grandes volumes consiste na extração da cor pelo aquecimento da uva entre 60 e 80 °C, separação do mosto e engarrafamento logo após pequeno descanso, para decantação das borras mais grossas. Os sucos obtidos por meio dessa tecnologia são turvos e geralmente apresentam depósitos de bitartarato de

potássio no fundo do recipiente. Outro processo consiste na maceração sulfurosa (adição de SO<sub>2</sub>) da uva esmagada, por alguns dias, para extrair a cor. A seguir, é realizado a separação do mosto sulfitado e sua conservação é feita em grandes recipientes até a comercialização, quando é dissulfitado em aparelho especial e engarrafado. Embora mais simples, essa técnica é pouco utilizada. O terceiro processo é uma combinação dos dois anteriores, pois consiste em fazer a extração da cor pelo aquecimento da uva esmagada, separar o mosto e conservá-lo até a comercialização na forma de mosto sulfitado, quando é dissulfitado e engarrafado. O suco de uva reprocessado ou reconstituído é elaborado do suco concentrado, gerando suco límpido e sem depósitos<sup>57</sup>.

## 2.8 CONTROLE DE QUALIDADE DOS SUCOS DE UVA

A composição dos sucos de uva é avaliada de forma frequente por um conjunto de parâmetros físico-químicos por meio de métodos analíticos clássicos e instrumentais para controle de qualidade, sendo obrigatório para sua comercialização. Na Química Analítica, os métodos clássicos envolvem medidas de massa (gravimetria) e volume (volumetria), enquanto que os métodos instrumentais compreendem o uso de instrumento para avaliar medidas de uma propriedade física nas espécies químicas ou em reações (condutividade, absorção ou emissão de radiação, potencial de eletrodo, razão massa/carga, entre outros)<sup>68,69</sup>. Os critérios estabelecidos na lei vigente são necessários para comprovar a procedência e qualificar o suco de uva, além de averiguar eventuais falhas no processo agroindustrial de produção<sup>21</sup>.

A instrução normativa nº 14/2018 do MAPA é a mais recente norma destinada a uva e seus derivados. Entretanto, nenhum dos ensaios previstos na legislação brasileira é capaz de autenticar se o suco de uva integral é do tipo orgânico ou convencional, por esse motivo a relevância dos estudos envolvidos na presente tese. Os parâmetros físico-químicos analisados nos sucos de uva integral, adoçado, reconstituído e gaseificado devem cumprir limites (valores mínimo e máximo) determinados pela legislação, conforme apresentado na Tabela 5<sup>64</sup>. A seguir, são explicados, brevemente, os princípios de cada método usado no monitoramento da qualidade dos sucos de uva.

**Tabela 5.** Parâmetros físico-químicos avaliados no controle de qualidade dos sucos de uva.

Método	Valor Mínimo	Valor Máximo
Sólidos solúveis, °Brix, a 20 °C	14,0	-
Sólidos insolúveis, % v/v	-	5,0
Sorbitol, g/L	-	0,2
Acidez total, mEq/L (pH 8,2)	55	-
Acidez volátil, mEq/L	-	10
Álcool etílico, % v/v a 20 °C	-	< 0,5
Florizina	Ausência	
Corante artificial	Ausência	
Edulcorante	Ausência	

Fonte: Brasil, 2018<sup>64</sup>.

### 2.8.1 Sólidos Solúveis

A determinação de sólidos solúveis pode ser estimada pela medida de seu índice de refração por comparação com tabelas de referência. O índice de refração de uma substância pura é uma constante, mantidas as condições de temperatura e pressão. A presença de sólidos solúveis na água resulta numa alteração do índice de refração. É possível determinar a quantidade de soluto pelo conhecimento do índice de refração da solução aquosa. O método consiste em adicionar o suco, com o auxílio de um conta-gotas, no prisma do refratômetro do tipo Abbé, dispositivo caracterizado pelo fácil manuseio e que utiliza quantidades mínimas de amostras, projetado para medir o índice de refração. Assim, é realizada leituras dos valores de sólidos solúveis totais e da temperatura, sendo essas medidas corrigidas em função da temperatura<sup>21,70</sup>.

### 2.8.2 Sólidos Insolúveis

O suco de uva não deve exibir fragmentos sólidos como partes de sementes ou películas da casca. Entretanto, é permitido apresentar-se turvo pela presença de partículas da polpa. Ainda, a estabilização física do suco no processo de produção origina a precipitação mineral natural de cristais de bitartarato de potássio, o que garante a legitimidade do produto. O método baseia-se na determinação de matérias insolúveis em água, filtradas sob condições específicas. O suco é pesado, diluído em água e fervido, substituindo a água evaporada. Após, é realizada filtração a vácuo com água quente e os sólidos insolúveis são secos em dessecador e pesados<sup>21,70</sup>.

### **2.8.3 Acidez Total**

A determinação de acidez pode fornecer um dado valioso na apreciação do estado de conservação de um produto alimentício. Um processo de decomposição, seja por hidrólise, oxidação ou fermentação, altera quase sempre a concentração dos íons de hidrogênio. Os métodos de determinação da acidez podem ser os que avaliam a acidez titulável ou fornecem a concentração de íons de hidrogênio livres, por meio do pH. Este método baseia-se na titulação de neutralização dos ácidos com solução padronizada de álcali, com uso de indicador fenolftaleína para soluções claras de vinho e outras bebidas alcoólicas fermentadas ou com o peagâmetro para soluções escuras. A acidez total corresponde à soma dos ácidos tituláveis quando se neutraliza o mosto ou suco de uva a pH 7,0 com solução alcalina. Portanto, a acidez total corresponde a soma da acidez fixa, no qual os ácidos tartárico e málico são os mais importantes, com a acidez volátil. Por princípio, quanto mais elevada for a acidez fixa, mais baixa é a volátil<sup>21,70</sup>.

### **2.8.4 Acidez Volátil**

A acidez volátil corresponde à soma dos ácidos graxos presentes no mosto ou no suco de uva no estado livre ou salificado. Está associada com a qualidade da uva com relação a sua sanidade e as etapas de produção. O suco de uva apresentará teor elevado de acidez volátil se a uva apresentar percentual alto de amadurecimento, sofreu esmagamento e compactação durante a colheita e transporte e não foi processada de imediato. O princípio do método é a separação dos ácidos voláteis por meio do arraste pelo vapor d'água e posterior titulação ácido-base do destilado com solução de hidróxido de sódio padronizada e uso de indicador fenolftaleína<sup>21,70</sup>.

### **2.8.5 Sorbitol, Florizina, Edulcorantes, Álcool Etílico e Corante Artificial**

A identificação e determinação de sorbitol, florizina e edulcorantes é realizada por meio da técnica analítica instrumental de Cromatografia Líquida de Alta Eficiência (CLAE) com padrões de referência<sup>70-72</sup>. O sorbitol é um glicídio e a sua quantificação pode ser bastante útil na detecção de adulterações. Quantidades superiores de sorbitol indicam mistura de sucos ou fermentados de maçã, pera e cereja as quais são ricas nesse composto<sup>73</sup>. A florizina é um composto fenólico que tem sido amplamente utilizado para

identificar adulteração de sucos de frutas e outros produtos, sendo o principal glicosídeo fenólico encontrado em maçãs<sup>72</sup>. Os edulcorantes são substâncias diferentes dos açúcares que conferem sabor doce aos alimentos. Seu uso justifica-se nos produtos destinados a consumidores que necessitam de restrição calórica em suas dietas, bem como para aqueles portadores de diabetes. Aspartame, acesulfame-K, ciclamato de sódio e sacarina são alguns exemplos de edulcorantes<sup>70</sup>.

O teor de álcool etílico no suco de uva permite avaliar a qualidade da uva utilizada no processamento e a adequação do processo de suquificação, podendo ser estipulado pelo uso da técnica analítica instrumental de Cromatografia Gasosa (CG). Um teor elevado de etanol é indício de que a uva sofreu um início de processo fermentativo<sup>21</sup>. Por fim, os corantes artificiais são aditivos alimentares que tem a função de tornar os alimentos coloridos e mais atraentes para consumo e nos sucos de uva integral a adição é proibida. O método aplicado baseia-se na separação dos corantes por cromatografia ascendente em papel e comparação dos fatores de retenção com padrões de corantes<sup>70</sup>.

## 2.9 TÉCNICAS DE ESPECTROMETRIA MOLECULAR

A espectroscopia pode ser definida como a ciência que estuda a interação dos diferentes tipos de radiação eletromagnética com a matéria. Os fenômenos espectroscópicos são avaliados por meio da aplicação prática de técnicas conhecidas pela terminologia espectrometria. As técnicas espectroscópicas, portanto, são classificadas de acordo com a região do espectro eletromagnético envolvida na medida. Os fundamentos teóricos das técnicas usadas nesta tese são abordados a seguir, dentre os quais a espectrometria de absorção na região do ultravioleta-visível, espectrometria de absorção na região do infravermelho e a espectrometria de fluorescência molecular.

### 2.9.1 Espectrometria na Região do Ultravioleta-Visível

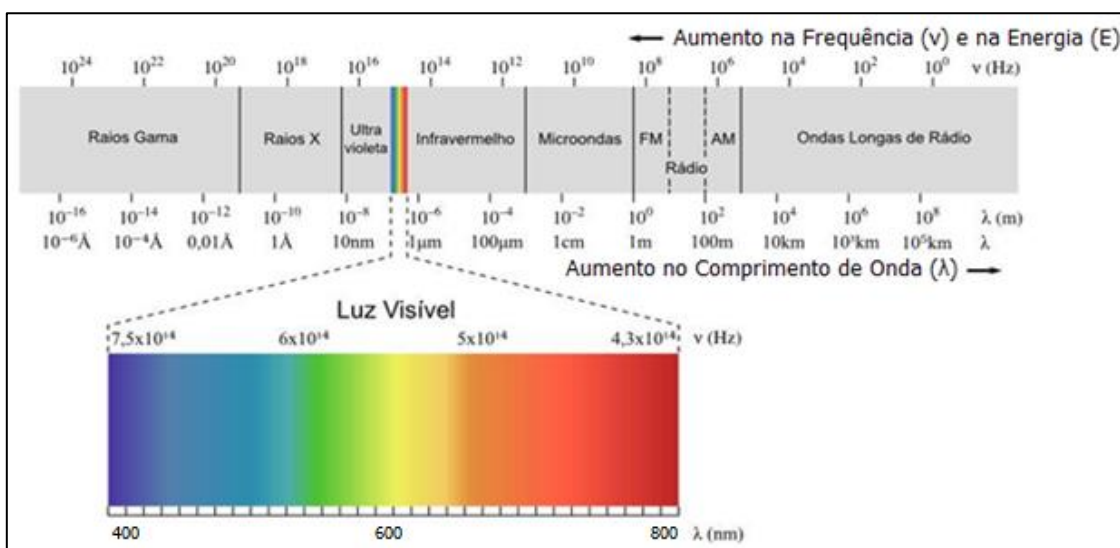
A espectrometria na região ultravioleta-visível (UV-Vis) é uma das técnicas analíticas mais utilizadas em diversas áreas da ciência para a análise quantitativa e qualitativa de muitas espécies inorgânicas, orgânicas e bioquímicas que absorvem radiação ultravioleta. A técnica é fundamentada nas medidas de absorção das moléculas estimuladas a realizar transições eletrônicas em função da absorção de energia



quantizada na região do UV-Vis. Em outras palavras, um espectrofotômetro UV-Vis mede a quantidade de luz absorvida por um determinado material em cada comprimento de onda ( $\lambda$ ) das regiões do ultravioleta e do visível no espectro eletromagnético<sup>68,74</sup>.

A absorção da radiação ultravioleta e visível é resultado da excitação de elétrons em moléculas com grupos cromóforos, no qual os comprimentos de onda dos picos de absorção podem ser correlacionados com os tipos de ligações das espécies moleculares, o que torna essa técnica útil também como uma ferramenta adicional para a identificação de grupos funcionais<sup>75</sup>. A luz ultravioleta é a radiação eletromagnética com comprimentos de onda na faixa entre 180 e 400 nm, enquanto que a luz visível tem comprimentos de onda no intervalo entre 400 e 800 nm<sup>68,74</sup>. A Figura 9 mostra o espectro eletromagnético com os tipos de radiação e os intervalos de comprimento de onda e de frequência.

**Figura 9.** Espectro eletromagnético com os tipos de radiação e os intervalos de comprimentos de onda e de frequência.



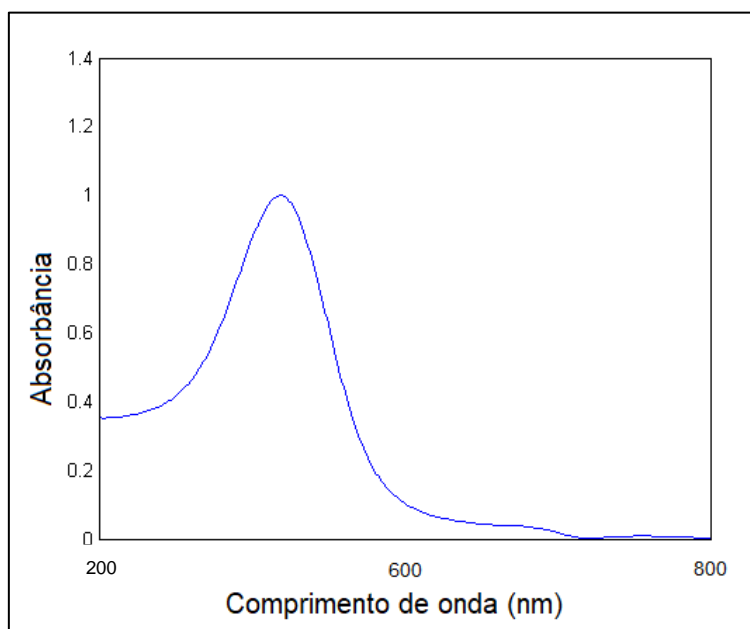
Fonte: adaptada de Ronan, 2007<sup>76</sup>.

A instrumentação usa um feixe contínuo de luz branca incidente sobre uma amostra que absorve luz e a radiação resultante emergente será detectável pela cor complementar da radiação absorvida<sup>68,74</sup>. A absorção óptica na região do ultravioleta-visível resulta da passagem de elétrons de orbitais no estado fundamental de energia mais baixa para orbitais de um estado excitado energético maior, que irá depender da estrutura eletrônica da molécula. Esse processo pode ser verificado por meio de um espectro

expresso em um gráfico de intensidade de absorção em função do comprimento de onda, conforme a Figura 10<sup>68,74,77-79</sup>.

A absorção eletrônica se sobrepõe a subníveis rotacionais e vibracionais e tem como consequência a formação de um espectro com aspecto de banda. Essa banda de absorção no espectro UV-Vis é caracterizada pela sua posição e intensidade. A posição relaciona-se ao comprimento de onda da radiação incidente com máxima absorção de energia, em que a energia é igual a necessária para que ocorra a transição eletrônica. A intensidade se refere a interação da energia incidente com a estrutura eletrônica do material que a recebe, de modo a permitir a passagem do elétron no estado fundamental para o estado excitado<sup>68</sup>.

**Figura 10.** Exemplo de espectro na região do ultravioleta-vível.



Fonte: elaborada pelo autor, 2022.

A estrutura eletrônica da molécula é modificada pela absorção de energia, isso ocorre devido às transições pela excitação de um elétron de um orbital molecular completamente ocupado (normalmente a um orbital  $\pi$  ligante) a um orbital de energia maior desocupado, que é geralmente o primeiro orbital antiligante ( $\sigma^*$  ou  $\pi^*$ ). A energia absorvida em uma transição eletrônica está relacionada com a frequência ( $\nu$ ) por meio da Equação 1,

$$\Delta E = h\nu = \frac{hc}{\lambda} \quad (1)$$

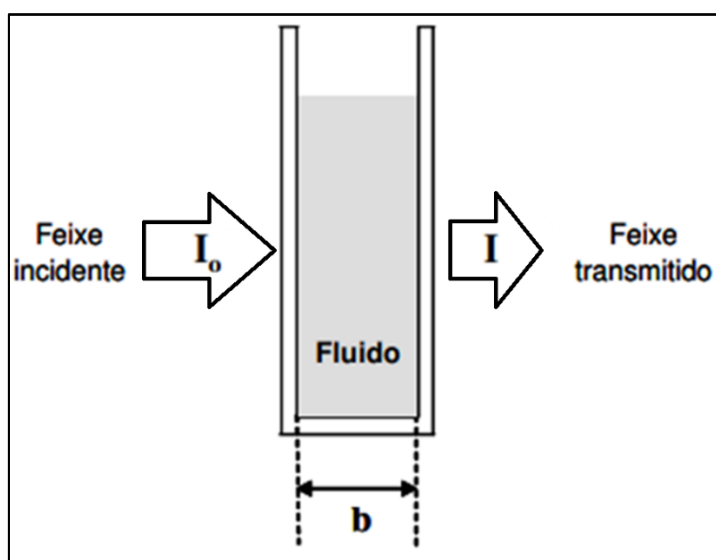
onde  $h$  é a constante de Planck,  $c$  é a velocidade da luz e  $\Delta E$  é a energia absorvida durante a transição, sendo a diferença de energia entre o estado fundamental e o estado excitado<sup>68,69</sup>.

A espectroscopia de absorção molecular se baseia na determinação da quantidade de luz que é absorvida pela amostra, descrito pela lei de Lambert-Beer, em que corresponde à diferença entre a intensidade da luz incidindo na solução ( $I_0$ ) e a intensidade da luz saindo da solução ( $I$ ), segundo a Figura 11. A quantidade de luz transmitida pode ser expressa tanto em transmitância como em absorbância de soluções contidas em uma célula (cubeta) transparente a radiação. A concentração de um analito absorvente se relaciona linearmente com a absorbância, conforme a Equação 2,

$$\left(\frac{I_0}{I}\right) = -\log T = A = \epsilon \cdot b \cdot c \quad (2)$$

onde  $A$  é a absorbância da solução,  $T$  é a transmitância da solução,  $\epsilon$  é a absorvidade molar (característico de cada substância),  $b$  é o caminho óptico (comprimento da célula onde está a amostra) e  $c$  é a concentração da solução<sup>68,69,74</sup>.

**Figura 11.** Representação do feixe incidente de luz sobre um fluido em uma cubeta com caminho óptico definido  $b$  e do feixe transmitido da radiação de luz.



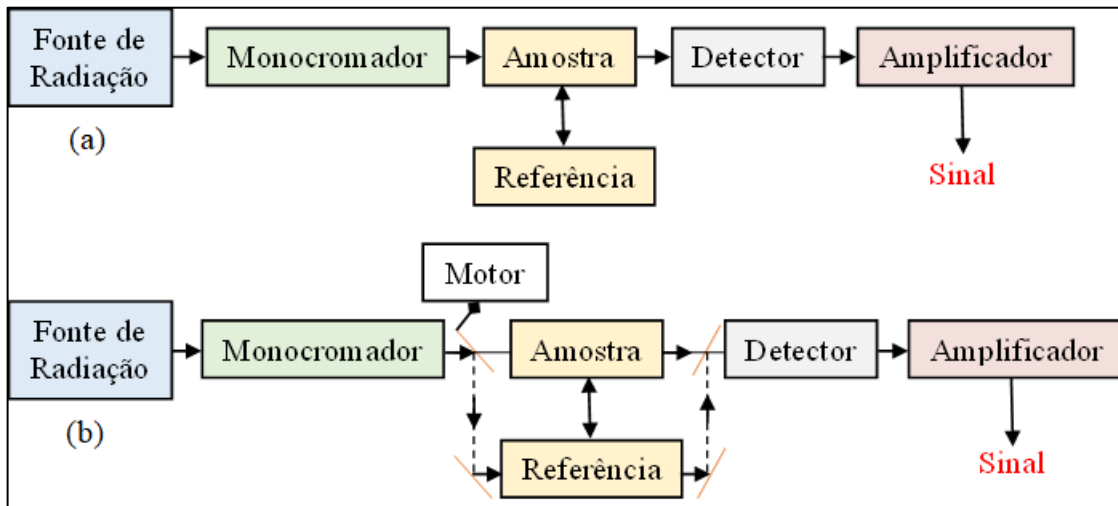
Fonte: adaptada de Harris, 2005<sup>69</sup>.

Os espectrofotômetros são instrumentos capazes de registrar dados de absorbância ou transmitância em função do comprimento de onda, em que o espectro de absorção é característico para cada espécie química. Em geral, os espectrofotômetros contêm cinco

componentes principais: fonte de radiação, monocromador, recipiente para conter a solução, detector e indicador de sinal. O equipamento pode ser de feixe simples, sendo necessário realizar a substituição da solução de referência (branco) pela amostra, ou de duplo feixe, em que dois feixes passam ao mesmo tempo pelos recipientes da referência e da amostra, conforme representado na Figura 12. Quando a região espectral usada é a ultravioleta, é necessário usar cubeta de quartzo para não absorver a radiação, enquanto que na região do visível é possível utilizar componente óptico de vidro<sup>68,69,75</sup>.

As vantagens do emprego do instrumento de feixe simples é que são opticamente mais simples e compactos e apresentam maior razão sinal/ ruído e, conseqüentemente, melhor limite de detecção. Por outro lado, os benefícios de se utilizar o dispositivo com duplo feixe se deve a rapidez e praticidade de operação, pois não necessita intercalar a amostra e a referência, e ser altamente reproduzível, no qual os efeitos eletrônicos e mecânicos nos feixes são iguais, ou seja, realiza uma correção automática para a perda de intensidade de luz à medida que o feixe passa pela amostra e pela solução de referência, uma vez que utiliza duas fontes monocromáticas, com comprimentos de onda diferentes<sup>68,69,75</sup>.

**Figura 12.** Representação de um espectrofotômetro de (a) feixe simples e (b) duplo feixe.



Fonte: adaptada de Holler, 2009<sup>68</sup>.

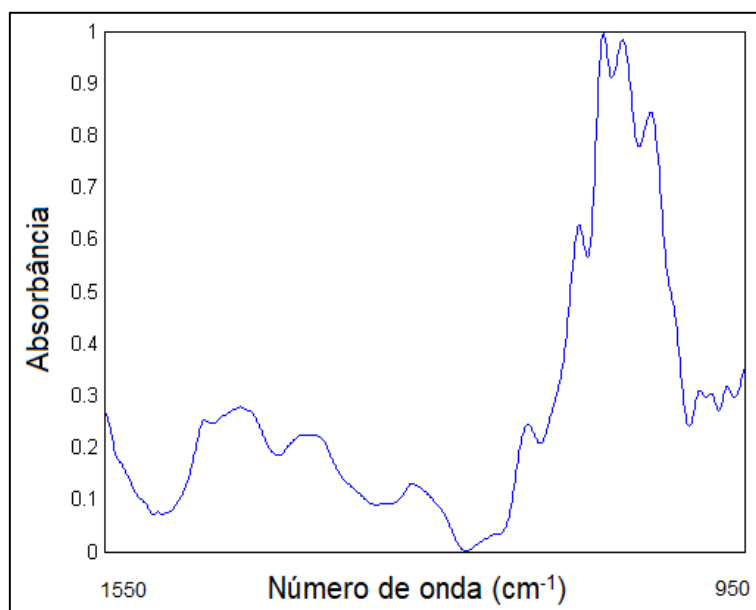
Portanto, a técnica da espectroscopia na região do UV-Vis é de grande aplicabilidade e destaca-se por apresentar como vantagens uma instrumentação de baixo custo, facilidade de operação, rapidez dos resultados, amostra não é destruída, exatidão nas medidas (erro relativo entre 1 a 5% em termos de concentração) e alta relação sinal-ruído (limites de detecção entre  $10^{-4}$  a  $10^{-5}$  mol.L<sup>-1</sup>)<sup>69,80</sup>.

## 2.9.2 Espectrometria na Região do Infravermelho

A espectrometria na região do infravermelho é uma técnica versátil aplicada às determinações qualitativas e quantitativas de espécies moleculares orgânicas e inorgânicas de todo tipo nos estados sólido, líquido e gasoso. A técnica baseia-se na absorção de energia quantizada pela interação das moléculas ou átomos com a radiação eletromagnética na região do infravermelho em um processo de vibração molecular, permitindo caracterizar propriedades dos compostos químicos. A absorção óptica na região do infravermelho é resultado de transições de baixa energia entre níveis vibracionais do estado fundamental para estado excitado vibracional no interior de um mesmo nível eletrônico<sup>68,77-79</sup>.

A frequência de uma absorção é consequência das massas dos átomos, das constantes de força das ligações e da geometria dos átomos na molécula. Assim, os valores das frequências de absorção da molécula no infravermelho são de suma importância, uma vez que permitem a interpretação do espectro, a fim de descrever os grupos funcionais presentes na estrutura química<sup>68,77-79</sup>. Logo, o espectro de infravermelho médio é composto por uma combinação de tons (frequências fundamentais), sobretons (*overtones* – harmônicas de frequência de pouca intensidade) e bandas de combinação de vibrações fundamentais dos grupos funcionais, conforme a Figura 13<sup>81</sup>.

**Figura 13.** Exemplo de espectro na região do infravermelho médio.



Fonte: elaborada pelo autor, 2022.

A radiação no infravermelho fica situada no espectro eletromagnético entre as regiões do micro-ondas e visível e é dividido habitualmente em três regiões espectrais: infravermelho próximo, médio e distante, conforme apresentado na Tabela 6<sup>68,75</sup>. A região espectral média no infravermelho é a de maior utilidade para a análise de funções orgânicas, pois cada grupo funcional absorve em um determinado número de onda<sup>68</sup>.

**Tabela 6.** Regiões espectrais no infravermelho.

Região espectral	Número de onda (cm <sup>-1</sup> )	Comprimento de onda (nm)	Frequência (Hz)
Próximo	12800 a 4000	780 a 2500	3,8 x 10 <sup>14</sup> a 1,2 x 10 <sup>14</sup>
Médio	4000 a 200	2500 a 50000	1,2 x 10 <sup>14</sup> a 6,0 x 10 <sup>12</sup>
Distante	200 a 10	50000 a 1000000	6,0 x 10 <sup>12</sup> a 3,0 x 10 <sup>11</sup>

Fonte: Holler, 2009<sup>68</sup>.

A região entre 4000 a 1400 cm<sup>-1</sup> é conhecida pela absorção da maioria dos grupos funcionais, enquanto que a região de 1400 a 200 cm<sup>-1</sup> é denominada de impressão digital, onde é possível identificar mais facilmente compostos, pois surge mudança notável dos picos de absorção pelas pequenas diferenças na estrutura da molécula<sup>75</sup>. A Tabela 7 apresenta de forma resumida as frequências de absorção no infravermelho de diversos grupos funcionais.

**Tabela 7.** Frequências de absorção no infravermelho de diversos grupos funcionais.

Ligação	Função Química	Frequência (cm <sup>-1</sup> )	Intensidade
C-H	Alcanos (estiramento)	3000-2850	Forte
	-CH <sub>3</sub> (deformação)	1450 e 1375	Médio
	-CH <sub>2</sub> - (deformação)	1465	Médio
	Alcenos (estiramento)	3100-3000	Médio
	(deformação fora plano)	1000-650	Forte
	Aromáticos (estiramento)	3150-3050	Forte
	(deformação fora plano)	900-690	Forte
	Alcinos (estiramento)	3300	Forte
		Aldeídos	2900-2800 2800-2700
C=C	Alcenos	1680-1600	Médio-Fraco
	Aromáticos	1600 e 1475	Médio-Fraco
C≡C	Alcinos	2250-2100	Médio-Fraco

C–O	Álcoois, Éteres, Ésteres, Ácidos Carboxílicos, Anidridos	1300-1000	Forte
C=O	Aldeídos	1740-1720	Forte
	Cetonas	1725-1705	Forte
	Ácidos Carboxílicos	1725-1700	Forte
	Ésteres	1750-1730	Forte
	Amidas	1680-1630	Forte
	Anidridos	1810 e 1760	Forte
C–N	Aminas	1350-1000	Médio-Forte
C=N	Iminas e Oximas	1690-1640	Fraco-Forte
C≡N	Nitrilas	2260-2240	Médio
O–H	Álcoois e Fenóis, livres	3650-3600	Médio
	Álcoois, H-ligado	3400-3200	Médio
	Ácidos Carboxílicos	3400-2400	Médio
N–H	Aminas e Amidas Primárias e Secundárias (estiramento)	3500-3100	Médio
	(deformação)	1640-1550	Médio-Forte
X=C=Y	Alenos, Cetonas, Isocianatos, Isotiocianatos	2270-1940	Médio-Forte
N–O	Nitro (R–NO <sub>2</sub> )	1550 e 1350	Forte
S–H	Mercaptanas	2550	Fraco
	Sulfóxidos	1050	Forte
S=O	Sulfonas, Sulfatos, Sulfonamidas	1375-1300 e 1350-1140	Forte
C–X	Fluoreto	1400-1000	Forte
	Cloreto	785-540	Forte
	Brometo, Iodeto	< 667	Forte

Fonte: adaptada de Holler, 2009<sup>68</sup> e Silverstein, 2007<sup>75</sup>.

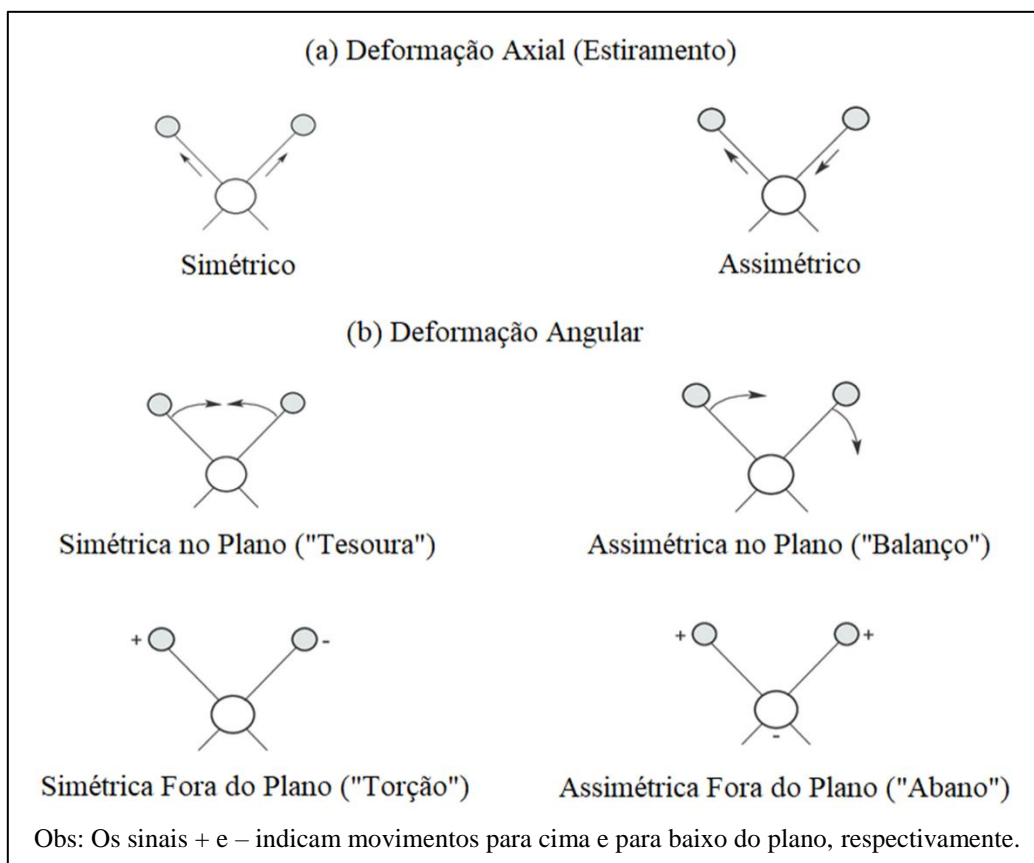
Os átomos na molécula oscilam continuamente em movimentos axiais e angulares nas ligações covalentes e a posição delas revelam os tipos de vibrações moleculares existentes<sup>68</sup>. A molécula deve apresentar variação no momento dipolo ( $\Delta\mu \neq 0$ ) durante seu movimento rotacional ou vibracional para absorver radiação no infravermelho. Dessa forma, o campo elétrico alternado da radiação pode interagir com a molécula e causar variação na amplitude de um de seus movimentos vibracionais ao redor das ligações covalentes nos átomos. O momento dipolar resultante ( $\mu_r$ ) depende da eletronegatividade dos elementos participantes da ligação e da geometria molecular<sup>68,75</sup>.

O espectro vibracional aparece como uma série de bandas denominadas de vibração-rotação, visto que a cada transição de nível de energia vibracional acontecem uma série de mudanças nos subníveis rotacionais e, deste modo, as linhas se

sobreposição<sup>75,77,78</sup>. As vibrações moleculares podem ser de deformação axial ou estiramento, que ocorre ao longo do eixo da ligação entre dois átomos e envolve uma variação constante na distância interatômica, ou de deformação angular, que são definidos por uma mudança no ângulo entre duas ligações, que são as vibrações ritmadas de ligações que tem um átomo em comum<sup>68,75</sup>.

As deformações axiais podem ser do tipo simétrica ou assimétrica e as deformações angulares podem ser de quatro tipos: simétrica ao plano, assimétrica ao plano, simétrica fora do plano ou assimétrica fora do plano. As moléculas não-lineares com N átomos tem  $3N - 6$  modos normais de vibração, enquanto que as moléculas lineares tem  $3N - 5$  devido a não observação da rotação em torno do eixo molecular. Os modos de vibrações moleculares de deformação axial e angular estão representados na Figura 14<sup>68,75</sup>.

**Figura 14.** Modos de vibrações moleculares de deformação axial e angular.



Fonte: adaptada de Holler, 2009<sup>68</sup>.

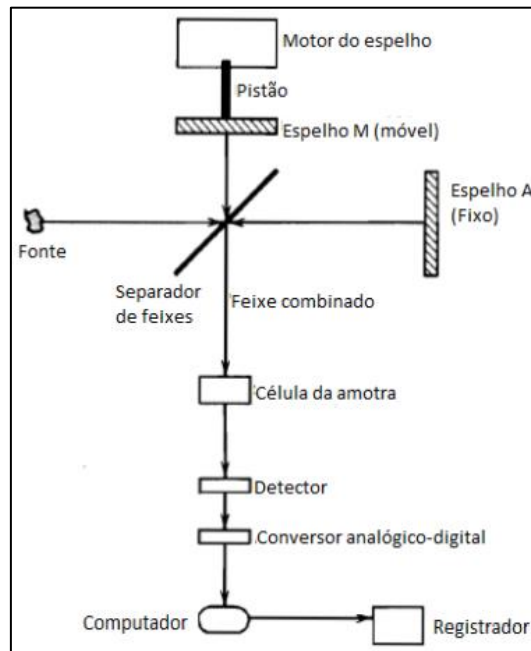
Os primeiros instrumentos no infravermelho são conhecidos como dispersivos, em que a varredura do espectro era feita com luz monocromática por meio de feixe simples ou duplo, apresentando baixa sensibilidade e demora na análise<sup>82</sup>. O



desenvolvimento tecnológico possibilitou a criação de equipamentos modernos classificados como não dispersivos com uso de transformada de Fourier, que é uma técnica matemática de separação de formas de ondas com diferentes frequências que se sobrepõem<sup>68,83</sup>. Esses espectrofotômetros no infravermelho com transformada de Fourier (FTIR) utilizam um interferômetro ao invés de um monocromador, o que propicia que todos os comprimentos de onda sejam detectados ao mesmo tempo<sup>75,82</sup>.

O FTIR opera por meio do uso de um jogo de espelhos, um estático e o outro móvel, em que é possível variar a distância percorrida por dois feixes de luz ocasionando uma série de interferências construtivas e destrutivas e, conseqüentemente, mudanças na intensidade da radiação recebida pelo detector, conhecido como interferograma. O interferograma é submetido à transformada de Fourier, que converte para a forma de frequência. Portanto, a amostra é submetida a uma larga faixa de energia que acarreta na vibração das moléculas cada qual com suas características, originando o espectro na região do infravermelho médio<sup>75,77,78,82</sup>. A Figura 15 mostra uma representação dos componentes de um espectrofotômetro de FTIR.

**Figura 15.** Representação de um espectrofotômetro de FTIR.



Fonte: Silverstein, 2007<sup>75</sup>.

A técnica FTIR possui diversas vantagens sobre os espectrômetros dispersivos como a rapidez de obtenção dos resultados, a simplicidade de operação, requer mínimo de preparo das amostras, a quantidade de analito é mínima, a amostra não é destruída na

análise, não é poluente, não gera resíduos tóxicos, apresenta elevada sensibilidade e seletividade, possui alta resolução espectral, é reprodutível, tem possibilidade de automação e o equipamento é de baixo custo quando comparado a outras ferramentas instrumentais, sendo facilmente encontrado em laboratórios. Além disso, é possível determinar a quantidade de componentes em uma mistura e identificar materiais desconhecidos<sup>35,68,83</sup>.

### 2.9.3 Espectrometria de Fluorescência Molecular

As técnicas de luminescência molecular se fundamentam em moléculas do analito que foram excitadas e que, ao retornarem ao seu estado fundamental, resultam na emissão de radiação eletromagnética na região do ultravioleta próximo do visível gerando, assim, um espectro com dados para análise qualitativa e quantitativa desses compostos. Quando a excitação das moléculas é em virtude da absorção de fótons, responsável pela transferência de elétrons de valência de um orbital menos energético para um orbital de maior energia, temos o fenômeno da fotoluminescência<sup>68,74,84</sup>.

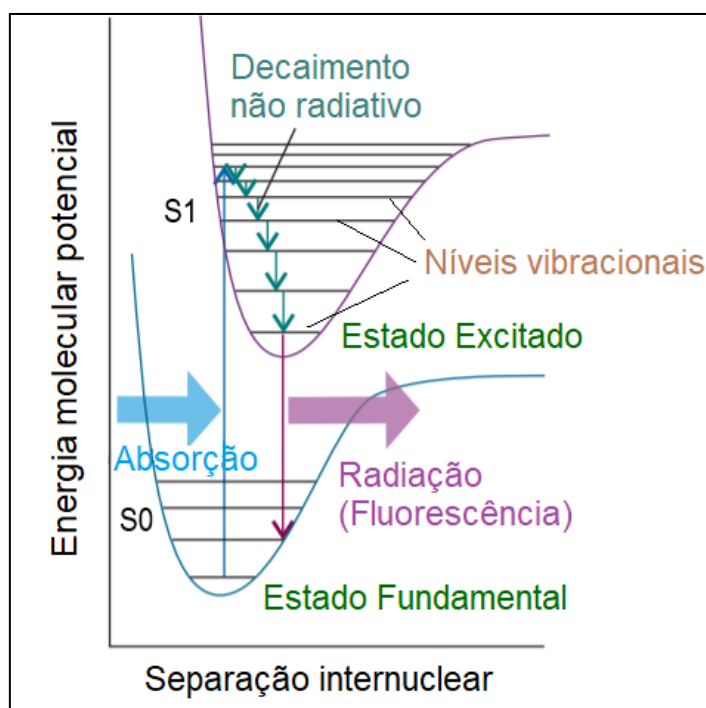
A ocorrência da fluorescência acontece quando o spin do elétron permanece com sua orientação original no estado excitado singleto e o regresso dessa população para o estado fundamental, que também apresenta caráter singleto, é permitido, ocorrendo muito rapidamente em nanosegundos. Desse modo, a fluorescência é a ocorrência luminescente mais frequente, visível facilmente na temperatura ambiente e diretamente em solução, e com metodologia experimental bastante acessível<sup>84-86</sup>.

A exibição da fluorescência ocorre respeitando as regras de seleção, que limitam o número de transições, segundo a Figura 16, conhecida como Diagrama Jablonski. As moléculas são estimuladas do estado fundamental ( $S_0$ ) pela absorção de radiação com comprimento de onda característico e promovidas para um estado excitado singleto ( $S_1$ ). Após a população de moléculas excitadas se desativam mediante uma série de relaxamentos (transições não radiativas) dos níveis vibracionais de estados eletrônicos de mesma multiplicidade até atingir o primeiro nível vibracional do estado excitado singleto de menor energia. Esse processo de relaxação é um evento que acontece muito rapidamente ( $10^{-13}$  a  $10^{-11}$  segundos) e sem emissão de radiação<sup>68,74,84,85</sup>.

A partir desse nível vibracional do estado excitado de menor energia, se a multiplicidade não alterar, as moléculas podem prosseguir dois percursos para retornar ao estado fundamental. Caso a separação de energia entre  $S_1$  e  $S_0$  não seja muito elevada

e haver possibilidade de sobreposição de níveis vibracionais, a população passa inicialmente para níveis vibracionais de S0 que possuem energia mais elevada (conhecido como cruzamento interno) sem emissão de radiação eletromagnética com posterior relaxamento vibracional até o estado vibracional fundamental de menor energia. Entretanto, se a diferença energética entre S1 e S0 for relativamente grande e as condições do meio favoráveis, a desativação para o estado fundamental se dá com emissão de radiação na forma de fluorescência<sup>68,74,84,85</sup>.

**Figura 16.** Diagrama Jablonski com as transições de energia do fenômeno da fluorescência.



Fonte: adaptada de Atkins, 2006<sup>84</sup>.

A fluorescência é a base da fluorimetria que engloba o conjunto de técnicas analíticas baseadas na detecção dos fótons emitidos por moléculas excitadas de caráter singleto quando estas retornam para o estado fundamental<sup>68</sup>. As substâncias que frequentemente exibem esse fenômeno são os compostos aromáticos, alifáticos ou alicíclicos contendo carbonila ou duplas ligações altamente conjugadas. A maioria dos hidrocarbonetos aromáticos não-substituídos fluorescem em solução e a eficiência quântica geralmente aumenta com o número de anéis e seu grau de condensação<sup>74,84,85</sup>.

A fluorescência mais intensa e mais útil é encontrada em compostos contendo grupos aromáticos com transições de baixa energia. As variáveis que afetam principalmente a fluorescência são a estrutura molecular e o ambiente químico. Estes

fatores também determinam a intensidade de emissão e quando a luminescência irá ocorrer. Isto posto, o controle de pH é necessário em medidas de fluorescência<sup>74,84,85</sup>.

## 2.10 FERRAMENTAS QUIMIOMÉTRICAS

A Quimiometria é uma área da química que utiliza métodos matemáticos e estatísticos com o propósito de planejar ou escolher melhores procedimentos de medidas, interpretar experimentos e obter o máximo de informações relevantes que podem ser extraídas da análise de um conjunto de dados de origem química. O vocábulo surgiu na década de 1970, na Química Analítica, como consequência do extenso desenvolvimento no campo da instrumentação, que começou a fornecer gradativamente mais dados, e nos microcomputadores, que permitiram processar essas informações. Os métodos quimiométricos podem ser utilizados em dados multivariados com as finalidades de planejamento experimental, análise exploratória, reconhecimento de padrões, classificação e calibração<sup>28,35,87,8</sup>.

### 2.10.1 Técnicas de Classificação Multivariada

As técnicas de análise multivariada têm se destacado, recentemente, graças ao desenvolvimento tecnológico, uma vez que é possível avaliar em um sistema muitas variáveis de uma amostra de forma simultânea, podendo, deste jeito, modelar informações físicas e químicas a partir dos dados espectrais, ou seja, transformar uma resposta instrumental no dado químico de interesse<sup>28,35,87</sup>.

Esses instrumentos de exploração multivariada demandam a aplicação de técnicas de álgebra linear, estatística multivariada e análise numérica, no qual quase todos os métodos quimiométricos são montados em termos de matrizes. A análise multivariada pode ser considerada conforme a configuração matricial em duas grandes classes: a análise exploratória e o reconhecimento de padrões, em que os dados são representados em uma única matriz, e os problemas de classificação e de calibração, quando se procura correlacionar duas ou mais matrizes<sup>28,35,87</sup>. De uma forma geral, as técnicas podem ser classificadas como não supervisionados e supervisionados.

Dentre as técnicas multivariadas não supervisionadas, em que não é considerado ou se conhece a que grupos pertencem os objetos de estudo, a análise de componentes principais (PCA) e a análise de agrupamentos hierárquicos (HCA) estão entre as

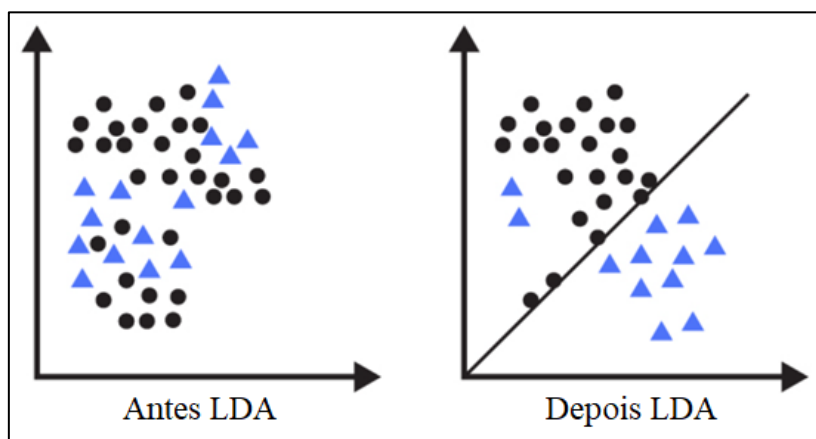
principais e mais notáveis, sendo gradativamente mais utilizadas com a finalidade de identificação de classes e reconhecimento da autenticidade e/ou adulteração de amostras com os dados químicos obtidos<sup>28,35,87,89</sup>.

Já entre as técnicas multivariadas chamadas de supervisionadas, onde se tem conhecimento prévio a que classes pertencem os objetos e essa informação disponível é aplicada na construção do modelo, a análise discriminante linear (LDA) e a análise discriminante por mínimos quadrados parciais (PLS-DA) estão entre as mais empregadas nos problemas de classificação em Química Analítica<sup>28,35,87,89</sup>.

### 2.10.1.1 Análise Discriminante Linear (LDA)

A análise discriminante linear (LDA, do inglês *Linear Discriminant Analysis*) foi desenvolvida por Ronald Fisher em 1936 para resolver a discriminação entre classes em problema de taxonomia<sup>90</sup>. Hoje é uma das técnicas de aprendizado de máquina mais populares e bem estabelecidas, aplicada em diversos campos do conhecimento para classificação e reconhecimento de padrões, conforme ilustrado na Figura 17. É uma técnica supervisionada que tem a finalidade de remover informações redundantes e extrair as características mais relevantes do conjunto de dados<sup>91,92</sup>.

**Figura 17.** Classificação de vários objetos antes e depois da implementação da LDA.



Fonte: adaptada de Mohammadi, 2015<sup>93</sup>.

O conceito da ferramenta é a necessidade de realizar uma redução de dimensionalidade dos dados a fim de simplificar o problema, por meio de algoritmos de seleção de variáveis, usando agrupamentos de classes conhecidas, expressa pela Equação 3. A LDA calcula combinações lineares entre duas ou mais funções discriminantes

(variáveis latentes) das variáveis originais medidas e não correlacionadas para caracterizar e prever a classe das observações fornecidas<sup>94,95</sup>.

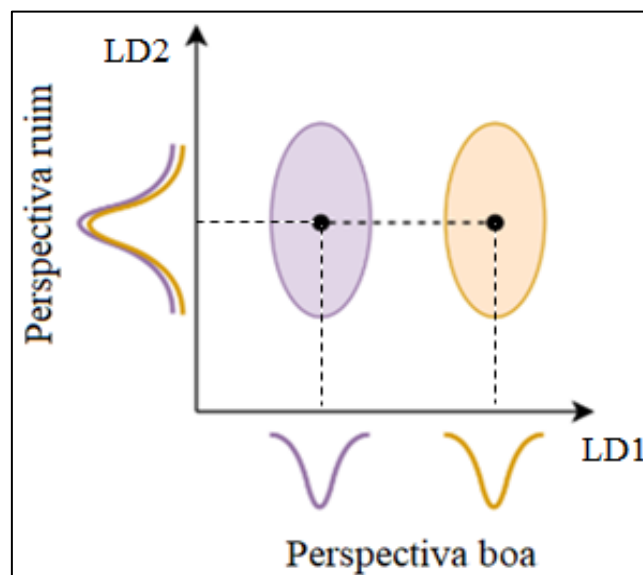
$$Z = \beta_1 \bar{X}_1 + \beta_2 \bar{X}_2 \quad (3)$$

onde Z são as combinações lineares, X é a média das variáveis das classes e os coeficientes  $\beta$  são os pesos das variáveis. Os coeficientes  $\beta$  são função da matriz de covariância C e do centro da média de cada classe envolvida  $\bar{X}_1$  e  $\bar{X}_2$ , apresentado na Equação 4 pela função discriminante linear representado pela distância de Mahalanobis<sup>95,96</sup>.

$$D_{1-2} = \sqrt{(\bar{X}_1 - \bar{X}_2)^T C^{-1} (\bar{X}_1 - \bar{X}_2)} \quad (4)$$

A técnica calcula os escores discriminantes para cada observação para classificar em que classe de variável de resposta está. Esses escores são obtidos pela localização de combinações lineares das variáveis independentes. É atribuído que as observações dentro de cada classe são extraídas de uma distribuição Gaussiana multivariada. O algoritmo projeta o conjunto de dados no espaço dimensional e seleciona a direção que maximiza os eixos componentes, chamados discriminantes lineares, para a separação entre as várias classes, conforme a Figura 18<sup>92,95,97</sup>.

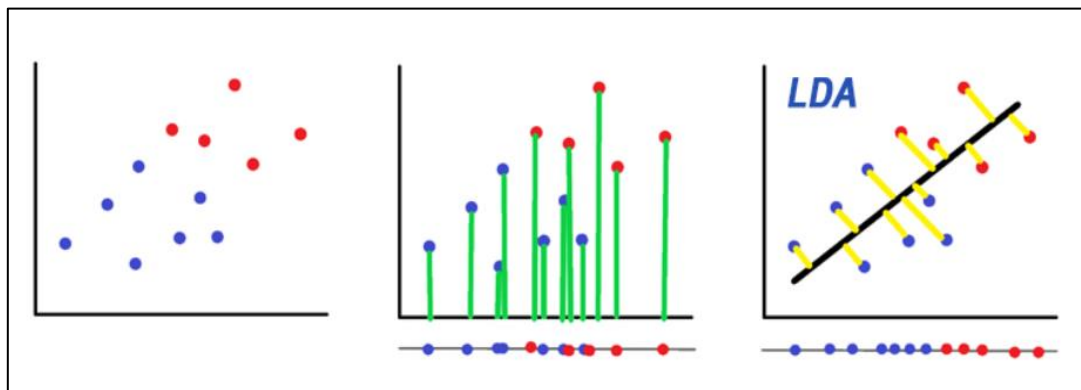
**Figura 18.** Aplicação da técnica LDA com a maximização dos eixos componentes para a separação das classes.



Fonte: adaptada de Ali, 2019<sup>97</sup>.

A LDA utiliza duas matrizes de dispersão diferentes, uma que calcula a variância entre as classes e outra que computa a variância dos padrões em relação a classe à qual pertence<sup>92,94,95</sup>. A discriminação das amostras nas classes é feita determinando o conjunto ótimo de pesos para as variáveis independentes do melhor conjunto de variáveis. A técnica cria um novo eixo e projeta os dados ortogonalmente de forma a minimizar a dispersão intraclasse e maximizar a separação entre classes distintas, conforme a Figura 19<sup>98,99</sup>. As limitações dos modelos LDA são a quantidade de variáveis que precisa ser menor que a quantidade de amostras envolvidas e as variáveis em análise não podem se correlacionar (colinearidade). A solução para essas restrições é o uso de algoritmos de seleção de variáveis<sup>91,100,101</sup>.

**Figura 19.** Projeção dos dados ortogonalmente pela técnica LDA.



Fonte: Silva, 2020<sup>99</sup>.

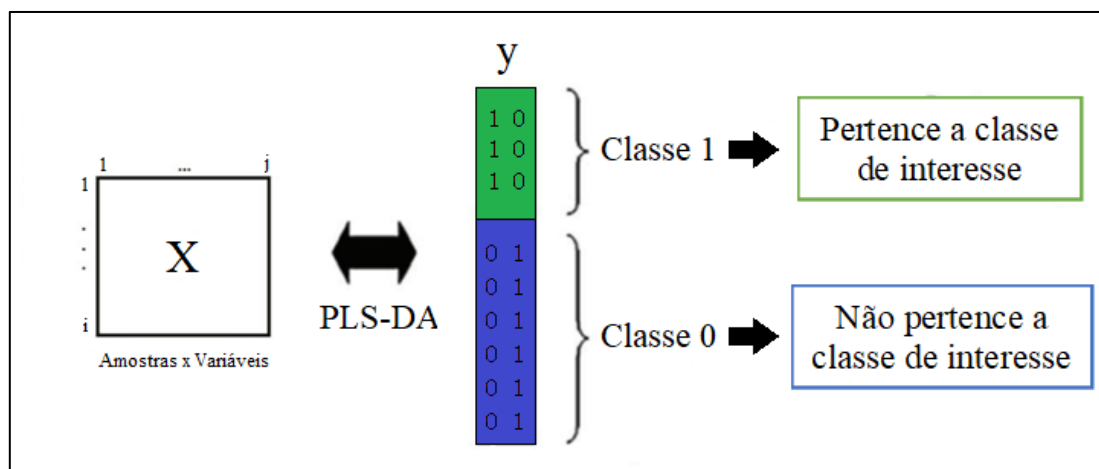
#### 2.10.1.2 Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA)

A análise discriminante por mínimos quadrados parciais (PLS-DA, do inglês *Partial Least Squares Discriminating Analysis*) é um modelo desenvolvido a partir da regressão por mínimos quadrados parciais (PLS), técnica popular para calibração multivariada. A PLS-DA é essencialmente a ferramenta PLS com uma variação na representação e na abordagem das variáveis dependentes. A técnica de classificação supervisionada PLS-DA se baseia no princípio de que muitas classes podem obter a separação máxima pela rotatividade dos componentes principais. É considerada uma modelagem descritiva, pois investiga a relação entre as variáveis do conjunto de dados, e preditiva, já que identifica padrões dentro do conjunto de dados e prevê comportamentos futuros<sup>28,102</sup>.

A técnica relaciona as variáveis originais da resposta instrumental (matriz independente X) em combinações lineares (variáveis latentes) com máximo de

covariância (independência entre variáveis aleatórias) entre as variáveis dependentes da propriedade de interesse (matriz Y). A propriedade de interesse é uma variável categórica que descreve a atribuição de classe da amostra. A discriminação de cada classe é definida na matriz Y pelas colunas contendo valores de 1 e 0, no qual o valor 1 é atribuído às amostras pertencentes à classe de interesse e o valor 0 é atribuído às demais amostras da outra classe<sup>103</sup>. A Figura 20 mostra a organização dos dados usando PLS-DA.

**Figura 20.** Organização dos dados para a construção do modelo de classificação usando PLS-DA.



Fonte: adaptada de Santana, 2020<sup>89</sup>.

O processo geral da análise consiste nas etapas de treinamento e validação. Na etapa de treinamento, são selecionadas as amostras para estabelecer uma relação entre as matrizes X e Y por meio da modelagem pelo algoritmo PLS. Na etapa de validação, um outro conjunto de amostras, que não faz parte da etapa de treinamento, é selecionado para ser testado e avaliar a eficiência do modelo. Assim, o modelo pode ser utilizado para discriminar novas amostras daquele mesmo material, geralmente nas mesmas condições das amostras utilizadas para a etapa de treinamento<sup>96</sup>. A essência da PLS-DA é diminuir o volume de dados originais X e trocá-los pela matriz de escores e pesos, potencializando a covariância entre X e Y<sup>104</sup>. A decomposição dos dados é realizada pelas Equações 5 e 6,

$$X = \sum_{a=1}^A t_a p_a^T + E \quad (5)$$



$$Y = \sum_{a=1}^A t_a q_a^T + F \quad (6)$$

onde  $t_a$  é o vetor de escores,  $p_a$  e  $q_a$  são os pesos referentes à primeira variável latente e  $E$  e  $F$  são as matrizes de erros de  $X$  e  $Y$ , respectivamente<sup>96</sup>.

A abordagem PLS-DA, portanto, é usar uma regressão que opera uma decomposição linear nas matrizes  $X$  e  $Y$  para calcular os parâmetros do modelo<sup>103</sup>. O limite para uma amostra pertencer a uma classe é fundamentado na suposição de distribuição normal e previsto pela probabilidade condicional bayesiana, técnica que usa o teorema de Bayes (formulação matemática) para calcular a probabilidade de um evento ocorrer, dado que outro evento ocorreu<sup>102,105</sup>. O desempenho de classificação do modelo está vinculado a escolha correta do número de variáveis latentes (VL) na construção do modelo PLS-DA. Número insuficiente de VL, há falta de ajuste no modelo, ou seja, não é utilizada toda a informação útil para sua construção. Do mesmo jeito, número elevado de VL, tem-se sobreajuste no modelo com adição de informações redundantes ou indesejáveis, como ruído espectral<sup>89,102,105</sup>.

O número de variáveis latentes é escolhido com base na validação cruzada, em que uma fração das amostras do conjunto de calibração é separada e utilizada para validação interna. O modelo de treinamento é construído usando diferentes VL e são previstas as amostras de validação interna para cada um dos modelos construídos. Os erros obtidos são armazenados e o processo é repetido até que todas as amostras de treinamento sejam previstas<sup>89,102,105</sup>.

Há duas variantes da metodologia, a PLS1-DA e a PLS2-DA, no qual estão associadas com a dimensionalidade do arranjo binário (0 ou 1) das propriedades modeladas. Na PLS1-DA, cada coluna de  $y$  é modelada individualmente. Se existirem duas classes (A e B), são construídos dois modelos. O primeiro modelo usa valores 1 para a classe A e 0 para as demais, e o segundo modelo utiliza valores de 1 para a classe B e 0 para as demais. O resultado é um conjunto de escores e variáveis latentes específico por variável em avaliação. Já na PLS2-DA, todas as colunas da matriz  $y$  são calculadas em um único conjunto de escores e pesos, o que provoca à restrição de empregar o mesmo número de variáveis latentes para modelar todas as classes<sup>28,89,102</sup>.

## 2.10.2 Técnicas de Seleção de Variáveis

Na análise multivariada, o conjunto de dados apresenta ampla dimensão com muitas variáveis, o que pode afetar desfavoravelmente o desempenho e a utilização dos métodos de regressão e de reconhecimento de padrões<sup>106,107</sup>. Na aplicação da técnica LDA em conjunto de dados espectrais, faz-se necessário reduzir o conjunto de variáveis espectrais. Um método eficiente de seleção de variáveis consiste em selecionar um pequeno subconjunto de variáveis, que contém as informações mais relevantes e úteis dos dados, capazes de construir modelos com melhor desempenho e minimizar o aspecto da multicolinearidade (variáveis independentes que possuem relações lineares diretas)<sup>108</sup>.

A seleção de variáveis é um processo de otimização que reduz esforço computacional e tempo de máquina, melhora a capacidade preditiva e gera modelos mais parcimoniosos, ou seja, com maior capacidade de generalização para prever amostras desconhecidas. É um mecanismo relevante quando se tem características não informativas, repetitivas e irrelevantes que não proporcionam aproveitamento útil<sup>106-108</sup>. Em razão do grande volume de dados complexos gerados pelas técnicas instrumentais, o desenvolvimento de algoritmos para a seleção de variáveis torna-se essencial para se obter uma melhora nos modelos de classificação<sup>106,109</sup>.

Os modelos matemáticos podem ser divididos em probabilísticos e determinísticos. Os modelos probabilísticos (ou estocásticos), como os algoritmos genético (GA) e otimização por colônia de formigas (ACO), incluem uma certa taxa de aleatoriedade a fim de fornecer uma distribuição de resultados possíveis (geram múltiplos subconjuntos de variáveis), ou seja, os modelos descrevem todos os resultados e fornecem uma medida de quão provável é que cada um ocorra. Além disso, demandam mais tempo de máquina e podem “escapar” do problema de mínimos locais (não escolha do melhor subconjunto de variáveis)<sup>110,111</sup>.

Em vez disso, os modelos determinísticos, como os algoritmos stepwise (SW) e projeções sucessivas (SPA), são dependentes dos parâmetros e das condições iniciais (*inputs*) escolhidos e visam produzir uma única solução que descreva o resultado do "experimento" (geram sempre o mesmo subconjunto de variáveis). A construção dos modelos exige menos esforço computacional do que os estocásticos, gerando resultados mais rápidos, entretanto, podem ficar “presos” num mínimo local<sup>110,111</sup>.

### 2.10.2.1 Algoritmo Genético (GA)

Dentre os métodos mais comumente empregados na área da Quimiometria, destaca-se o algoritmo genético (GA, do inglês *Genetic Algorithm*). É uma ferramenta matemática de otimização meta-heurística de sistemas complexos inventada por John Holland em 1975<sup>112</sup>. Meta-heurística é um recurso para explorar espaços de busca usando diferentes estratégias. O GA simula o mecanismo de seleção genética natural inspirado no princípio da evolução biológica proposta por Charles Darwin em 1859. A teoria de Darwin propõe que o cromossomo biológico contém genes responsáveis pelas características próprias para cada indivíduo, no qual os indivíduos mais adaptados sobrevivem e transmitem seus atributos para as gerações seguintes<sup>113-115</sup>.

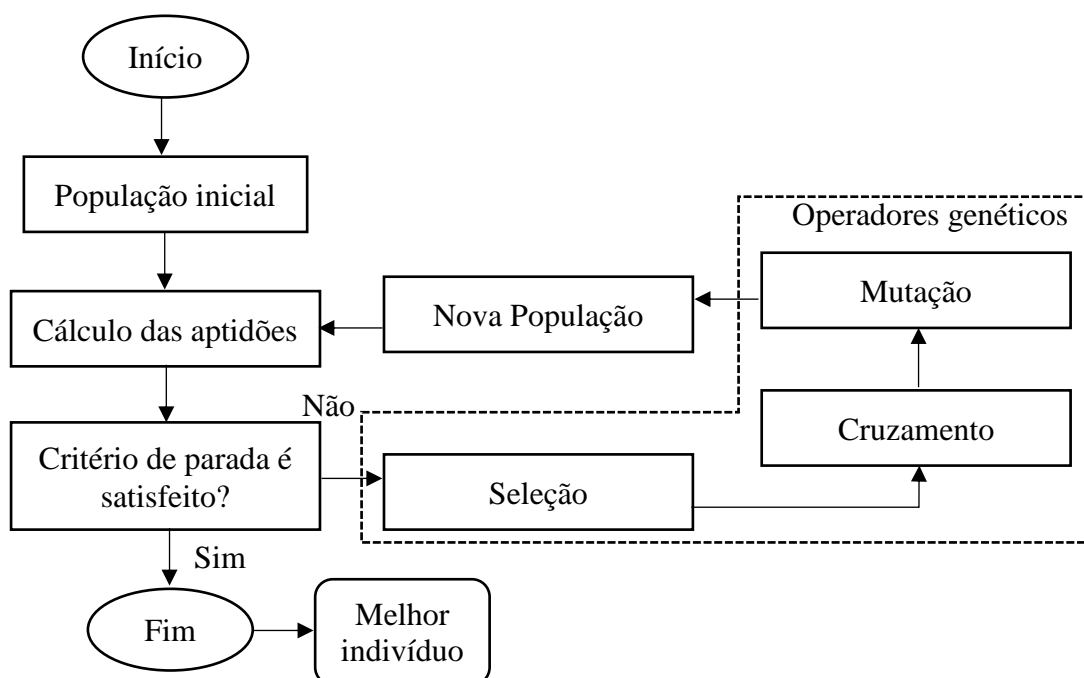
Desse modo, o algoritmo genético é estabelecido em simular computacionalmente um processo de adaptação natural com o desenvolvimento de um sistema artificial. O método fundamenta-se pela busca de soluções com base em um sistema evolutivo, em que o conjunto inicial é iterativamente melhorado e as gerações derivadas são mais evoluídas do que os antecedentes com os indivíduos mais frágeis sendo eliminados. O método GA mostra-se útil para seleção de variáveis a fim de eliminar a correlação aleatória e multicolinearidade entre as variáveis com as vantagens de auto-organização e auto-adaptação<sup>110,116</sup>.

O algoritmo genético é estruturado nas seguintes etapas: geração da população inicial, avaliação da população, seleção, recombinação, mutação e reinserção. Essas etapas são repetidas até que um critério de término seja alcançado. A nova geração substitui a anterior após a reprodução e mutação e o algoritmo continua a partir da avaliação da resposta. Cada indivíduo corresponde a uma solução possível e as variáveis são como “genes” que formam um “cromossomo”, que será testado quanto à sua aptidão, ou seja, a capacidade de fornecer a melhor classificação, também denominada como função objetivo<sup>110,114</sup>.

Inicialmente o GA gera uma população inicial, onde cada indivíduo é composto por um grupo de genes (variáveis) que corresponde a uma possível solução. A seguir, a população é avaliada com relação a aptidão (função objetivo do problema), se o indivíduo está apto a reproduzir e permanecer na população. É definido um critério de término de investigação do algoritmo como número de gerações, tempo do processo ou resultado obtido. A sequência é a seleção dos indivíduos com melhores cromossomos com propósito de garantir a geração de indivíduos melhores<sup>117-120</sup>.

A recombinação/cruzamento é a etapa de reprodução dos indivíduos com troca de genes e a mutação é uma modificação aleatória no gene de um indivíduo. Esses passos alteram os genes dos cromossomos da população ao longo das gerações, promovendo a diversidade e também a preservação dos aspectos dominantes e significativos dos antecessores, com geração de nova população. Por fim, as etapas são refeitas até atender a parada, em que o “melhor” conjunto de indivíduos (variáveis) são selecionados<sup>117-120</sup>. A Figura 21 mostra um fluxograma com as etapas envolvidas do algoritmo genético.

**Figura 21.** Fluxograma com as etapas envolvidas do Algoritmo Genético.



Fonte: elaborada pelo autor, 2022.

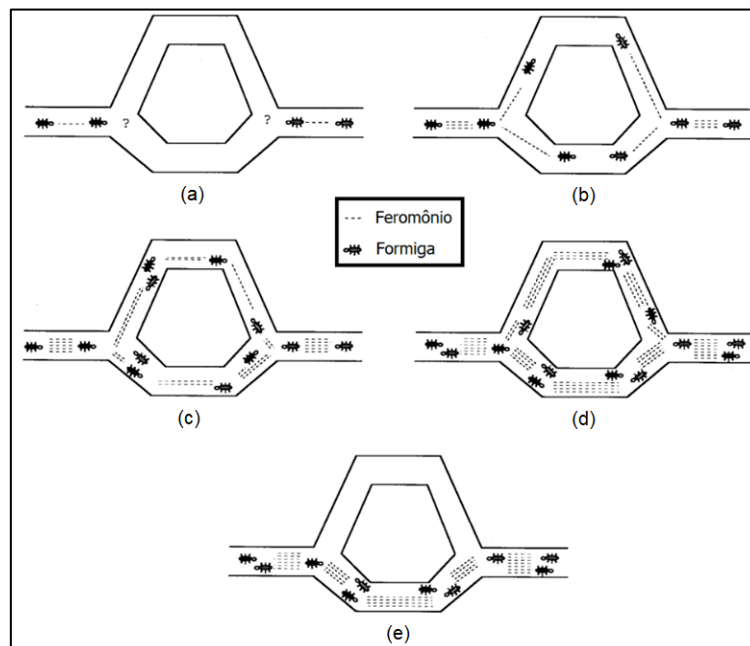
Quando se aborda a seleção de variáveis em processos químicos como nas técnicas espectroscópicas, o GA codifica cada gene no cromossomo como uma das variáveis do espectro e que precisa ser otimizado. Logo, o cromossomo contém todas as variáveis do espectro. A codificação binária é utilizada, onde cada gene pode ser selecionado (valor 1) ou não inserido (valor 0) no modelo, resultando em um subconjunto de variáveis escolhido<sup>117-119</sup>. As vantagens do algoritmo genético em relação a outras técnicas de otimização são a simples execução, aplicável a uma grande variedade de problemas, usa regras de transição probabilísticas e não determinísticas, apresenta maior flexibilidade no processamento do problema a ser resolvido, entre outras<sup>120</sup>.

### 2.10.2.2 Algoritmo de Otimização por Colônia de Formigas (ACO)

Outro método estocástico frequentemente empregado é o algoritmo de otimização por colônia de formigas (ACO, do inglês *Ant Colony Optimisation*), que corresponde a uma técnica de inteligência artificial desenvolvida por Dorigo e Gambardela em 1997 para resolução do problema do caixeiro viajante. O modelo matemático foi baseado na disseminação de formigas que deveriam percorrer, uma única vez, um conjunto de cidades na menor distância possível<sup>121,122</sup>. O método corresponde a classe de meta-heurísticos, isto é, um algoritmo de alto nível customizável para solucionar processos complicados de otimização que utilizam combinação de escolhas aleatórias e conhecimento histórico dos resultados anteriores para se guiarem<sup>123,124</sup>.

O algoritmo ACO é inspirado no comportamento coletivo das formigas com elevado grau de coordenação na procura do trajeto mais curto entre seu ninho e as fontes de alimento<sup>123,124</sup>. A Figura 22 mostra a sequência ((a), (b), (c), (d) e (e)) de comportamento das formigas explicado pela presença de feromônios nos trajetos. No início as formigas exploram de modo aleatório o ambiente ao redor de seu formigueiro de um jeito aparentemente desorganizado conforme procuram alimento. Ao moverem-se neste caminho, liberam compostos químicos conhecidos como feromônios, que evaporam com o tempo<sup>123-125</sup>.

**Figura 22.** Comportamento sequencial das formigas explicado pela presença de feromônios nos trajetos.



Fonte: adaptada de Dorigo, 1997<sup>126</sup>.

As formigas são induzidas pela presença dessas substâncias biologicamente ativas no ambiente e apresentam uma tendência probabilística a trilhar no caminho em que a concentração de feromônios é mais forte. As formigas que transitam o menor trajeto, retornam ao formigueiro mais rapidamente. Desse modo, a trilha percorrida possui maior concentração de feromônio. As novas formigas, que se localizam em bifurcações, vão ter preferência ao caminho com maior quantidade da substância. Portanto, a escolha das formigas é fundamentada na quantidade de feromônios acumulada na trilha das formigas antecedentes e na distância em cada trajeto, a fim de otimizar a distância percorrida entre o ninho e a fonte de alimento<sup>123-125</sup>.

A quantidade de feromônios é atribuído um valor de peso diferente por meio dos parâmetros  $\alpha$  (alfa) e  $\beta$  (beta), que servem para determinar a influência da substância e da informação heurística. A atratividade de cada caminho pode ser expressa pela Equação 7<sup>124,126</sup>.

$$a_{ij}(t) = \frac{[\tau_{ij}(t)]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in N_i} [\tau_{il}(t)]^\alpha [\eta_{il}]^\beta} \quad (7)$$

A atratividade  $a_{ij}(t)$  é calculada para uma formiga que está em um nó  $i$  qualquer, e para todo nó  $j$  pertencente a  $N_i$ , em que  $N_i$  é o conjunto dos nós adjacentes ao nó  $i$ . A variável  $\tau_{ij}$  representa a quantidade de feromônio acumulada no trajeto  $ij$ , e a variável  $\eta_{ij}$  é igual ao inverso do comprimento do trajeto  $ij$ . O parâmetro  $t$  indica que os valores variam a cada iteração. Depois do cálculo de todos os valores dos caminhos percorridos, é calculado a probabilidade de uma formiga  $k$  escolher um caminho  $ij$ , conforme a Equação 8. O parâmetro  $N_i^k$  é o subconjunto de  $N_i$  que apresenta apenas os nós que ainda não foram visitados pela formiga  $k$ <sup>124,126</sup>.

$$p_{ij}^k(t) = \frac{a_{ij}(t)}{\sum_{l \in N_i^k} a_{il}(t)} \quad (8)$$

A formiga  $k$  seleciona um caminho entre todas as possíveis alternativas por meio do cálculo da probabilidade acumulada a partir das probabilidades de cada trajeto  $p_{ij}$  e sorteia-se um número aleatório no intervalo entre 0 e 1. O caminho escolhido será aquele em que o número sorteado possuir valor maior ou igual à  $p_{ij} - 1$  e menor que  $p_{ij}$ . Em

virtude de todas as formigas  $k$  percorreram os caminhos, o feromônio é atualizado em todos os trajetos conforme a evaporação e deposição<sup>124,126</sup>. A atualização dos feromônios é dada pela Equação 9.

$$\tau_{ij}(t) \leftarrow (1 - p)\tau_{ij}(t) + \sum_{k=1}^m \Delta\tau_{ij}^k(t) \quad (9)$$

O parâmetro  $p$  representa o nível de evaporação dos feromônios, que atua como um coeficiente de decaimento. O somatório representa todas as contribuições individuais de cada formiga  $k$  que tenha passado pelo caminho  $ij$  na atual iteração,  $m$  é o número de formigas e o  $\Delta\tau_{ij}^k$  é a quantidade de feromônios que a formiga  $k$  deposita no trajeto  $ij$ , e é definida como o inverso do comprimento total do percurso que a formiga  $k$  percorreu<sup>124,126</sup>.

Na abordagem de seleção de variáveis, o algoritmo caracteriza as variáveis como dimensões espaciais, definindo os possíveis caminhos a serem seguidos pelas formigas, com as coordenadas permitidas. A cada geração, as formigas depositam uma certa quantidade de feromônio, que aumenta de acordo com a função objetivo definida por cada caminho. A quantidade de feromônio acumulada em cada uma das coordenadas de dimensão ajuda a formiga a encontrar o melhor caminho. Dessa forma, um determinado caminho está relacionado a um conjunto de variáveis selecionadas, que por sua vez correspondem a uma determinada solução. O algoritmo otimiza os modelos com base em variáveis pré-definidas e trabalha no conceito de obrigar a acumulação de feromônios, para descartar variáveis irrelevantes e selecionar as “melhores” variáveis<sup>127,128</sup>.

### 2.10.2.3 Algoritmo Stepwise (SW)

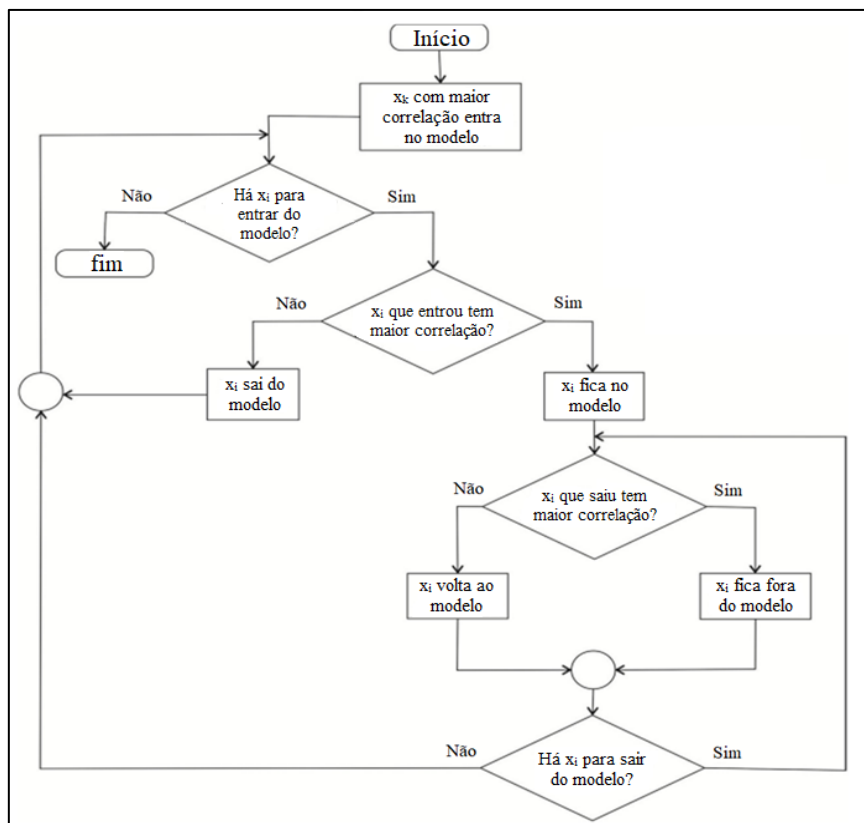
O algoritmo *Stepwise* é um método heurístico, em que é encontrado boas soluções na maioria das vezes, porém sem garantia de sempre obtê-la. O método se baseia nos conceitos de seleção de variáveis progressiva (Forward Selection) e eliminação (Backward Elimination)<sup>129</sup>. O modelo Forward Selection inicia pela escolha de uma variável independente  $x$  que possui a maior correlação com variável dependente  $y$ , e outras variáveis são adicionadas em seguida. O cálculo estatístico do teste F é aplicado para análise depois de cada variável adicionada, em que se determina a soma quadrática

residual do novo modelo para calcular um valor de  $F$  ( $F_{cal}$ ). A variável que apresenta o maior valor de  $F_{cal}$  permanece no modelo<sup>108</sup>.

Por outro lado, o modelo Backward Elimination começa com todas as variáveis disponíveis e vai removendo-as. O método realiza uma avaliação com base no teste  $F$  a cada saída de variável, em que a variável com menor valor de  $F_{cal}$  é retirada do modelo. Em ambas as metodologias o processo segue até que não haja variáveis com valores de  $F_{cal}$  maiores que um valor crítico tabelado ( $F_{crítico}$ ) ou obtido experimentalmente para um determinado nível de confiança e grau de liberdade<sup>130</sup>.

O algoritmo Stepwise busca maximizar a razão entre as medidas das dispersões dentro da classe (soma da distância quadrada entre uma amostra e a média de todas as amostras dentro da classe) e entre as classes (soma da distância quadrada entre as médias da classe e todas as amostras). O método é útil na presença de muitos preditores, selecionando automaticamente as “melhores” variáveis a serem usadas no modelo. Desta forma, as variáveis redundantes são removidas e são identificadas as variáveis que foram mais significativas para a função de discriminação<sup>131,132</sup>. A Figura 23 apresenta um fluxograma com as tomadas de decisão do método Stepwise<sup>133</sup>.

**Figura 23.** Fluxograma com as etapas envolvidas do Algoritmo Stepwise.



Fonte: Alves, 2013<sup>133</sup>.



O algoritmo inicia selecionando a variável  $x_k$  que resulta na maior razão entre a soma dos quadrados total e a soma dos quadrados dentro da categoria, e a próxima variável  $x_i$  com a maior razão entre as demais é selecionada na sequência, de forma a não ultrapassar um valor limite de correlação definido pelo usuário<sup>131,132</sup>. A ferramenta abrange uma fase de inclusão seguida por uma etapa de exclusão ponderada por meio de testes F parciais para as  $x$  variáveis independentes a fim de que cada variável seja reinserida no cálculo e assim reavaliar quando agrupada no modelo<sup>108</sup>.

O método foi desenvolvido para economizar esforços computacionais, pois não necessita realizar todas as regressões possíveis, atingindo um conjunto ótimo de variáveis independentes. As limitações da regressão Stepwise incluem considerar que exista apenas um único conjunto ótimo de variáveis independentes, o que frequentemente não ocorre, e quando as variáveis independentes são altamente correlacionadas, o que resulta em inferiores previsões<sup>134</sup>.

#### 2.10.2.4 Algoritmo das Projeções Sucessivas (SPA)

Um dos métodos determinísticos mais populares é o algoritmo das projeções sucessivas (SPA, do inglês *Successive Projections Algorithm*), desenvolvido com a intenção de selecionar variáveis com mínimo de correlação em calibração multivariada. É um método que utiliza os princípios da Forward Selection, em que inicia de uma variável independente  $x$  e se vai incluindo uma nova variável em cada iteração com menor multicolinearidade com respeito as já escolhidas<sup>129</sup>.

O método SPA compreende sobretudo três etapas para problemas de calibração. Na primeira, são aplicadas operações nas colunas da matriz de calibração para selecionar subconjuntos de variáveis pressupondo mínima multicolinearidade. O segundo passo é definido o subconjunto que apresenta melhor resultado em relação ao parâmetro *RMSEV* (Raiz Quadrada do Erro Médio Quadrático de Validação, do inglês *Root Mean Square Error of Validation*), de acordo com a Equação 10. Na última fase, o subconjunto escolhido é submetido a um procedimento de eliminação para determinar se alguma variável pode ser removida sem perda significativa da capacidade de previsão<sup>129,135,136</sup>.

$$RMSEV = \sqrt{\frac{1}{N_{val}} \sum_{i=1}^{N_{val}} (y_{val}^i - \hat{y}_{val}^i)^2} \quad (10)$$

onde  $N_{val}$  é o número de amostras do conjunto de validação e  $y_{val}^i$  e  $\hat{y}_{val}^i$  são os valores de referência e os previstos para o parâmetro de interesse nas amostras de validação, respectivamente. O *RMSEV* avalia a habilidade de previsão de um modelo<sup>129,135,136</sup>.

O algoritmo possui uma variante usada na modelagem de métodos de reconhecimento de padrões supervisionados para classificação, em que usa a função de risco médio  $G$  ao invés da função de custo *RMSEV*, conforme a Equação 11.

$$G = \frac{1}{N_{val}} \sum_{i=1}^{N_{val}} g_i \quad (11)$$

onde  $g_i$  é o risco de uma classificação incorreta da amostra de validação, sendo definido pela Equação 12<sup>137</sup>.

$$g_i = \frac{r^2[X_i, \bar{X}(I_i)]}{\min_{j \neq I_i} r^2[X_i, \bar{X}(I_j)]} \quad (12)$$

O numerador  $r^2[X_i, \bar{X}(I_i)]$  é o quadrado da distância de Mahalanobis entre o objeto  $X_i$  com índice de classe  $I_i$  e a média amostral  $\bar{X}(I_i)$  de sua classe verdadeira. O denominador corresponde ao quadrado da distância de Mahalanobis entre o objeto  $X_i$  e o centro da classe errada mais próxima. Um pequeno valor de  $g_i$  aponta que  $X_i$  está perto do centro da sua verdadeira classe e distante dos centros das demais classes. Portanto, minimizar a função de custo  $G$  resulta em melhor separação dos objetos<sup>138</sup>.

Em resumo, o SPA propõe que subconjuntos de variáveis sejam formados de acordo com uma sequência de operações de projeção, que envolvem as informações da matriz de respostas instrumentais. Quando aplicado em dados de técnica de espectrometria molecular, o algoritmo inicia com um comprimento de onda e é calculado o vetor de projeção no espaço ortogonal e, então, se incorpora uma nova variável a cada iteração com critério de parada definido pelo usuário, até que um número específico de comprimentos de onda seja selecionado<sup>138,139</sup>.

A seleção de variáveis é formulada como um problema de otimização combinatória restrita, no qual subconjuntos de variáveis são testados e comparados com relação ao desempenho do modelo resultante. Logo, a combinação de variáveis pode ser obtida com o mínimo de colinearidade e de informações redundantes, o que é útil para

melhorar a capacidade de previsão do modelo. A metodologia fornece uma solução única ao final do processo de seleção<sup>135,138</sup>.

### 2.10.3 Pré-Tratamento dos Dados

O pré-tratamento dos dados é uma etapa primordial de preparação do conjunto de dados experimentais organizados na forma de matriz e é realizada antes dos procedimentos quimiométricos com intuito de aprimorar a eficácia e a validade dos modelos multivariados. O objetivo é suprimir informações não relevantes por meio da extração matemática de fontes de variação inconvenientes, que não seriam retiradas espontaneamente no decorrer da análise, e do tratamento das matrizes de dados para tornar os cálculos mais adequados antes da modelagem<sup>28,35,85,140</sup>.

Portanto, o pré-tratamento consiste num conjunto de técnicas de mineração de dados que envolvem preparação, organização e estruturação dos dados para transformar dados brutos em formatos adequados e eficientes. A seleção da técnica e combinações de pré-tratamento de dados requer uma abordagem sistemática e depende do objetivo da análise e das características dos dados em questão, em que vários procedimentos são testados para garantir que o pré-tratamento mais adequado seja utilizado. Há inúmeras técnicas de pré-tratamento de dados como alisamento, normalização, autoescalamento, centralização na média, cálculo de primeira ou segunda derivada para correção da linha de base, entre outras<sup>28,35,85,140</sup>. A seguir, são explicados brevemente o propósito de algumas dessas técnicas.

A ferramenta de alisamento ou suavização tem por finalidade diminuir o ruído aleatório causado, por exemplo, pelas vibrações mecânicas do equipamento, movimento das moléculas, ar ou luz que interferem no sinal analítico, por meio de algoritmos matemáticos como o filtro Savitzky-Golay, um dos mais empregados em tratamento de dados. Ainda, tem como objetivo o aumento da razão sinal/ruído, visto que os sinais instrumentais são compostos pelo sinal verdadeiro e ruído aleatório<sup>28,35,85,141</sup>.

O recurso da normalização possui como escopo a redução da influência de oscilações indesejadas existentes no conjunto de dados e é realizada dividindo cada variável por uma constante, tendo como resultado que todas as amostras ficam numa mesma escala. A técnica de autoescalamento implica em subtrair de cada elemento de uma coluna da matriz de dados o valor médio da respectiva coluna e dividir o resultado

pelo desvio-padrão dessa coluna, com a finalidade de remover diferenças de escala entre as variáveis e garantir que cada variável contribua igualmente para a análise<sup>28,35,85,141</sup>.

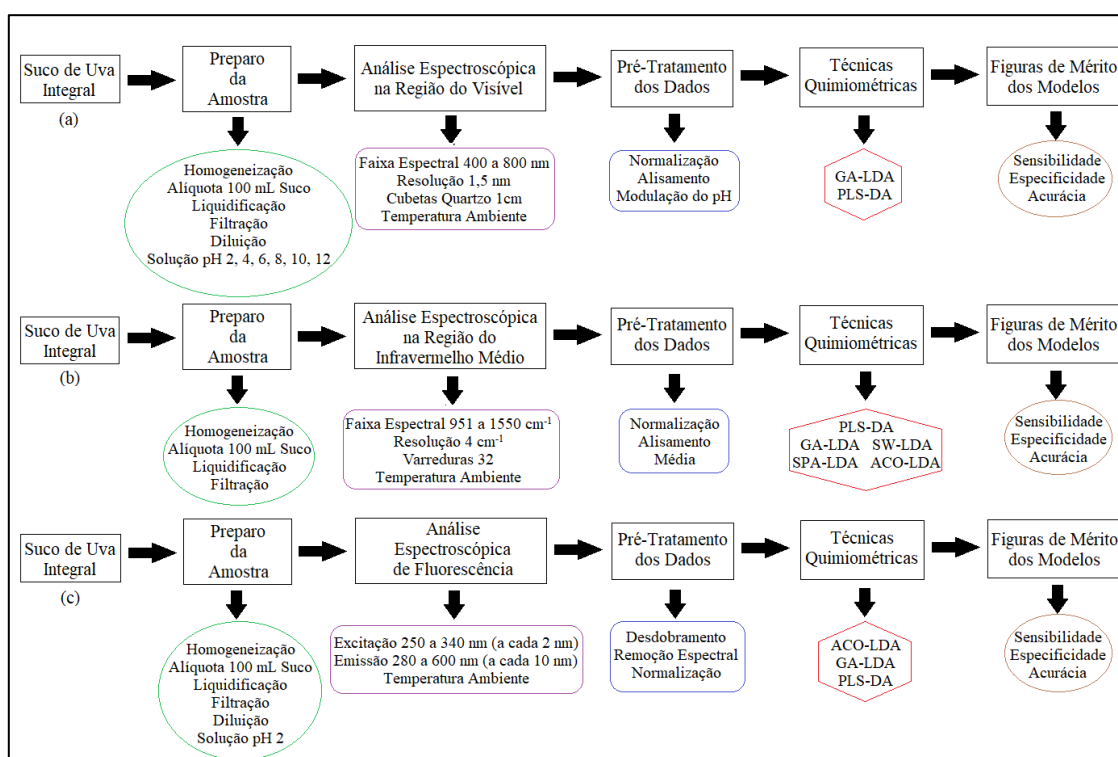
O intuito de centralizar os dados na média é de deslocar o sistema de eixos para o centro dos dados para que os pontos mais afastados não induzam mais do que os sinais mais perto do centro. A técnica consiste em primeiro se calcular o valor médio para cada variável e subtrair este valor para cada elemento de uma variável. Deste modo, cada variável passa a ter média zero e, então, o sistema de coordenadas é movido para o centro dos dados e as diferenças nas intensidades relativas das variáveis são melhores perceptíveis<sup>28,35,85,141</sup>.

Por fim, a técnica do cálculo da derivada tem como objetivo a correção do deslocamento da linha de base do espectro produzido para melhorar a separação de sinais sobrepostos. O algoritmo matemático mais utilizado para essa finalidade é o Savitzky-Golay, em que o cálculo da primeira e da segunda derivada acarretam alterações nas informações processadas, e essas modificações podem ter resultado positivo ao revelar um sinal analítico de interesse e negativo ao salientar ruídos instrumentais. Logo, o uso da primeira derivada corrige o espectro deslocado da linha de base, já a aplicação da segunda derivada conserta o espectro que apresenta um problema de inclinação na linha de base<sup>28,35,85,141</sup>.

### 3. METODOLOGIA EXPERIMENTAL

Todos os procedimentos metodológicos desenvolvidos na UFRGS nessa tese com as etapas elaboradas para cada estudo proposto podem ser visualizadas, resumidamente, em fluxograma, conforme a Figura 24. O item (a) se refere à pesquisa com dados espectrais na região do visível; o item (b) se refere ao trabalho com dados espectrais na região do infravermelho médio; e o item (c) se refere ao estudo com dados espectrais de fluorescência.

**Figura 24.** Fluxograma com as etapas envolvidas nos procedimentos metodológicos para cada estudo proposto.



Fonte: elaborada pelo autor, 2022.

#### 3.1 REAGENTES E MATERIAIS

No estudo com dados na região do visível, os seguintes reagentes químicos foram utilizados para obter as soluções tampão com diferentes valores de pH (2, 4, 6, 8, 10 e 12): carbonato de sódio ( $\text{Na}_2\text{CO}_3$ ), dihidrogenofosfato de potássio ( $\text{KH}_2\text{PO}_4$ ) e fosfato dipotássico ( $\text{K}_2\text{HPO}_4$ ) da Dinâmica, Indaiatuba, SP, Brasil; bicarbonato de sódio ( $\text{NaHCO}_3$ ) da Neon, Suzano, SP, Brasil; ácido acético glacial ( $\text{CH}_3\text{COOH}$ ) e acetato de

sódio anidro ( $\text{CH}_3\text{COONa}$ ) de Synth, Diadema, SP, Brasil; e ácido fosfórico ( $\text{H}_3\text{PO}_4$ ) de Anidrol, Diadema, SP, Brasil.

Para a investigação na espectroscopia de fluorescência, a solução tampão com valor de pH igual a 2 foi preparada com os seguintes reagentes químicos: dihidrogenofosfato de potássio ( $\text{KH}_2\text{PO}_4$ ) da Dinâmica, Indaiatuba, SP, Brasil, e ácido fosfórico ( $\text{H}_3\text{PO}_4$ ) da Anidrol, Diadema, SP, Brasil. No trabalho envolvendo os dados no infravermelho médio não foi necessário o uso de reagentes químicos.

### 3.2 AMOSTRAS DE SUCOS DE UVA

Todos os sucos de uva integral varietais (46 amostras obtidas de 11 variedades) foram produzidos nos meses de janeiro e fevereiro de 2019 pela Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Unidade Uva e Vinho de Bento Gonçalves, RS, Brasil, e disponibilizados para a pesquisa. Do total de 40 amostras de sucos de uva tinta, 28 amostras foram provenientes de um experimento que teve como objetivo avaliar variedades de videiras *Vitis labrusca* e híbridos interespecíficos criados e cultivados no Brasil quanto à capacidade de produzir uvas adequadas para a elaboração de sucos naturais em sistema de produção orgânica (sem o uso de fertilizantes químicos ou fungicidas). Este experimento avaliou as variáveis "castas de uvas tintas" (10 variedades, a saber: Bordô, Concord, Isabel, Isabel Precoce, BRS-Magna, BRS-Rúbea, BRS-Violeta, BRS-Cora, BRS-Carmem e Seleção 13), "porta-enxerto" (dois porta-enxertos, nomeadamente: 043-43 e Paulsen 1103) e "sistema de cultivo" (cultivo tradicional com as vinhas conduzidas em pérgola e cultivo em pérgola com as plantas protegidas por lona plástica).

Para efeitos de comparação, foram também utilizados 18 sucos de uva produzidos em videiras enxertadas em sistema de cultivo convencional ou não biológico (produção com recurso a fertilizantes químicos e fungicidas sintéticos para o controle de doenças fúngicas de folhas e bagos), sendo 12 amostras de castas tintas (nomeadamente: Bordô origem Minas Gerais, Bordô origem Muçum-RS, Bordô clone Grano d'Oro, Concord, Isabel, Isabel Precoce, BRS-Magna origem Muçum-RS, BRS-Magna, BRS-Rúbea, BRS-Carmem, BRS-Cora e BRS-Violeta) e 6 amostras de uva rosé da variedade Niágara rosada, produzidas de um experimento onde foram avaliados diferentes níveis de adubação das videiras. Finalmente, 44 sucos de uva integral foram obtidos de uvas cultivadas em vinhedos experimentais da EMBRAPA na cidade de Bento Gonçalves-RS, Brasil, e 2

sucos de uva (Bordô e BRS-Magna) foram obtidos com uvas de vinhedos localizados no município de Muçum-RS, distante 60 km da cidade de Bento Gonçalves.

O preparo do suco de uva integral em pequena escala foi feito por aquecimento e troca de calor em equipamento denominado suquificador integral (Figura 25), desenvolvido pela EMBRAPA em parceria com a empresa de refrigeradores de líquidos HBSR/Monofrio. O suquificador integral é feito de aço inoxidável, possui revestimento duplo contendo líquido de aquecimento em seu interior e construído com ângulo de inclinação. No seu interior, contém um tambor perfurado, onde as uvas são espremidas. O referido tambor gira em torno de um eixo central, facilitando a homogeneização da massa de fruta triturada. O equipamento é controlado por um sistema eletrônico que permite regular a temperatura e o tempo de aquecimento, bem como a velocidade e o regime de rotação do tambor interno. Possui ainda válvula de alívio de pressão e válvula inferior para retirada do suco<sup>142</sup>.

**Figura 25.** Suquificador integral para a produção de sucos de uva integral.



Fonte: Guerra, 2016<sup>142</sup>.

As uvas (50 kg por lote) foram previamente desengaçadas e esmagadas. O regime de homogeneização da uva esmagada no suquificador foi de um minuto de giro do tambor interno, intercalado com dois minutos de pausa. Nessas condições, a temperatura máxima do suco (70 °C) foi atingida 60 minutos após o início da suquificação. O líquido foi

drenado pela válvula inferior e a fase sólida foi prensada suavemente. Os dois sucos (obtidos por escorrimento e prensagem) foram misturados e resfriados a 0 °C, em câmara fria. Nessas condições, a borra foi completamente decantada após 12 horas, momento em que o suco foi separado da borra por sifonamento, pasteurizado a 75 °C no suquificador e engarrafado em recipientes de vidro de 1 litro de capacidade unitária. Os recipientes foram, então, resfriados e armazenados em uma adega climatizada entre 18 e 20 °C. O protocolo operacional foi padronizado para todos os sucos de uva<sup>142</sup>.

### 3.3 PREPARAÇÃO DAS AMOSTRAS

No Laboratório de Quimiometria e Instrumentação Analítica (LAQIA) da UFRGS, o suco de uva integral envasado e preparado pela EMBRAPA permaneceu armazenado em temperatura ambiente e livre da exposição solar até a sua análise. Assim, quando foi realizado o ensaio, o suco de uva integral foi, inicialmente, agitado manualmente, de forma suave, por 10 inversões na garrafa para homogeneização em temperatura ambiente. Em seguida, a garrafa foi aberta e uma alíquota de 100 mL de suco de uva foi coletada em um béquer, transferida para um liquidificador e mantida sob agitação por 1 minuto, a fim de obter mais uniformidade da amostra. O suco liquidificado foi submetido à filtração simples com papel de filtro qualitativo redondo 100 mm de gramatura 80 g/m<sup>2</sup> para remoção das partículas sólidas, com objetivo de não interferir na posterior análise espectroscópica. O suco filtrado foi colocado em tubo cônico do tipo Falcon. O trabalho relacionado à Química Analítica Verde usando dados na região do infravermelho médio envolveram até essas etapas.

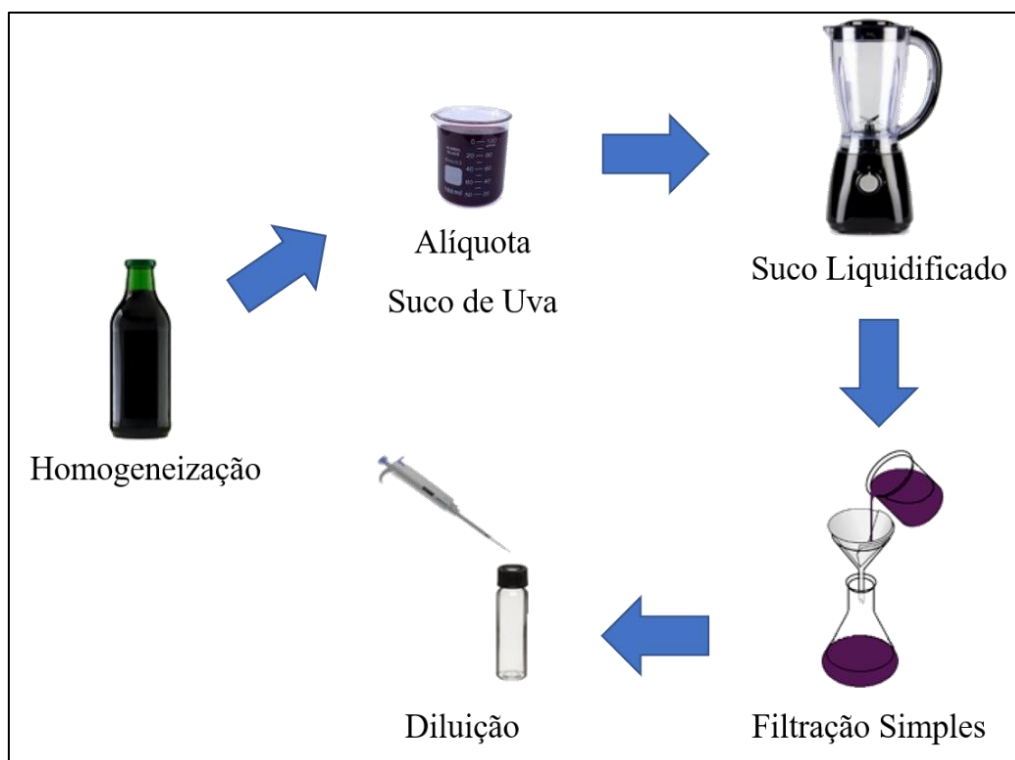
Para a pesquisa de espectroscopia na região do visível, além dos passos anteriores, as amostras foram preparadas diluindo 450 µL de suco de uva filtrado com 15 mL de solução tampão, a fim de adquirir soluções com diferentes valores de pH (2, 4, 6, 8, 10 e 12). Da mesma forma, para o estudo de espectroscopia de fluorescência, as amostras foram preparadas na mesma proporção de diluição, entretanto apenas elaborado solução com pH igual a 2. A escolha por essa razão de diluição de suco de uva com solução tampão deve-se a trabalho anterior<sup>143</sup>, em que os autores analisaram matéria-prima similar (vinho) em diferentes pHs. Ainda assim, foram avaliadas também diferentes proporções de diluição em sucos de uva, no entanto devido tanto a saturação quanto a perda de intensidade da banda espectral observada na posterior análise espectroscópica, foi definido esta como a melhor proporção de diluição. Todas as amostras, antes da análise



nos instrumentos de espectroscopia, foram avaliadas em peagâmetro portátil calibrado Metrohm<sup>®</sup> 827 pH lab (Metrohm AG, Suíça) a fim de constatar se o pH definido estava condizente.

A Figura 26 ilustra as etapas envolvidas na preparação das amostras de suco de uva integral. Na mesma data em que as amostras de suco de uva foram preparadas, igualmente foram realizadas as análises químicas para aquisição dos dados espectroscópicos (visível, infravermelho médio e fluorescência). Ressalta-se, ainda, que entre a elaboração do suco de uva integral na EMBRAPA e a preparação das amostras na UFRGS, o tempo decorrido foi de até 6 meses, a fim de que o suco não sofresse processo de oxidação e conseqüente degradação.

**Figura 26.** Etapas envolvidas na preparação das amostras de suco de uva integral.



Fonte: elaborada pelo autor, 2022.

### 3.4 MEDIDAS DE ESPECTROSCOPIA NA REGIÃO DO VISÍVEL

As amostras foram analisadas à temperatura ambiente em cubetas de quartzo com caminho óptico de 1 cm em espectrofotômetro UV-Vis. Para tanto, foi utilizado um espectrofotômetro Varian Cary<sup>®</sup> 50 UV-Vis (Agilent Technologies Inc., EUA) equipado com uma lâmpada de pulso de xenônio, monocromador Czerny-Turner, suporte de célula

de amostra frontal, detector de diodo de silício (Si) duplo e conectado a um microcomputador com Microsoft® Windows® (Microsoft Corporation, EUA). O equipamento empregado era de feixe duplo, em que dois feixes passam ao mesmo tempo pelos recipientes da solução de referência (branco - água destilada) e da amostra. As medições visíveis de absorção corresponderam à faixa espectral com comprimentos de onda de 400 a 800 nm e resolução de 1,5 nm. Os dados espectrais visíveis foram exportados em formato CSV.

### 3.5 MEDIDAS DE ESPECTROSCOPIA NA REGIÃO DO INFRAVERMELHO MÉDIO

Os espectros infravermelhos foram coletados em um espectrômetro Cary 630 FTIR (Agilent Technologies Inc., EUA) com um cristal de seleneto de zinco (ZnSe) e laser de hélio/neon. Todas as amostras foram analisadas em triplicata e a temperatura ambiente. Os espectros foram gerados em uma faixa entre  $951\text{ cm}^{-1}$  e  $1550\text{ cm}^{-1}$ , com resolução de  $4\text{ cm}^{-1}$  e 32 varreduras. Os dados espectrais de infravermelho foram exportados em formato ASP.

### 3.6 MEDIDAS DE ESPECTROSCOPIA DE FLUORESCÊNCIA

Os espectros de fluorescência foram obtidos usando um espectrofotômetro Varian Cary® Eclipse (Agilent Technologies Inc., EUA) equipado com uma lâmpada flash de xenônio, dois monocromadores Czerny-Turner de excitação e emissão, um suporte de célula de amostra frontal e conectado a um microcomputador por meio de uma interface serial RS232 (GPIB). Um detector PMT 600 V com o software Cary Eclipse foi usado para obter os dados e uma voltagem fotomultiplicadora média foi usada para evitar a saturação do detector. As larguras de fenda do monocromador no modo de excitação foram ajustadas para 10 nm e no modo de emissão ajustadas para 2 nm. Os espectros foram registrados nas faixas de comprimento de onda de 250 a 340 nm para o modo de excitação e 280 a 600 nm para o modo de emissão. Os dados de fluorescência espectral foram exportados em formato ASCII.

### 3.7 ANÁLISE QUIMIOMÉTRICA

Os cálculos quimiométricos foram realizados no ambiente MATLAB<sup>®</sup> versão R2010a 7.10.0.499 (The MathWorks Inc., EUA). Os modelos de classificação LDA e seleção de variáveis foram produzidos utilizando o VS TOOLBOX, uma caixa de ferramentas desenvolvida pelo grupo de pesquisa do LAQIA da UFRGS<sup>144</sup>. Os modelos PLS-DA foram desenvolvidos usando uma toolbox gratuita fornecida pela Milano Chemometrics<sup>102</sup>. A seleção de variáveis a partir do algoritmo genético (GA) foi utilizada em todos os estudos desenvolvidos na presente tese. O algoritmo de otimização por colônia de formigas (ACO) foi empregado nas técnicas espectroscópicas na região do infravermelho médio e de fluorescência. Por fim, os algoritmos Stepwise (SW) e de projeções sucessivas (SPA) foram usados apenas na pesquisa com dados FTIR.

Para a construção dos modelos LDA, o conjunto de amostras foi particionado pelo método de Kennard-Stone. O algoritmo seleciona o conjunto de amostras mais representativas e distribuídas homogêneas baseado nas distâncias Euclidianas entre elas, no qual as amostras mais representativas são utilizadas para o conjunto de treinamento e as amostras menos representativas são empregadas no conjunto teste<sup>145,146</sup>. Assim, 2/3 das amostras corresponderam ao conjunto de treinamento e 1/3 das amostras representaram ao conjunto de teste, proporção usada na literatura para a divisão de dados<sup>89</sup>. A divisão das amostras para os estudos de espectroscopia na região do visível e de fluorescência foi de 31 amostras selecionadas para o conjunto de treinamento (12 convencionais e 19 orgânicas) e 15 amostras para o conjunto de teste (6 convencionais e 9 orgânicas). Na pesquisa envolvendo a espectroscopia na região do infravermelho médio, 30 amostras foram selecionadas para o conjunto de treinamento (12 convencionais e 18 orgânicas) e 15 amostras foram selecionadas para o conjunto de teste (6 convencionais e 9 orgânicas).

### 3.8 VALIDAÇÃO ANALÍTICA DOS MÉTODOS

A validação de modelos de classificação multivariada pode ser realizada a partir do cálculo de parâmetros que assegurem que o modelo apresente desempenho adequado e dentro dos objetivos desejados. As figuras de mérito utilizadas para avaliar os modelos de classificação LDA e PLS-DA foram a sensibilidade, a especificidade e a acurácia, calculados pelas equações abaixo. A sensibilidade representa a proporção de positivos

identificados corretamente, ou seja, amostras pertencentes à classe e classificadas corretamente nesta classe (Equação 13); a especificidade/seletividade representa a proporção de negativos identificados corretamente, isto é, amostras não pertencentes à classe modelada e corretamente classificadas como não pertencentes (Equação 14); e a acurácia representa a proporção de amostras corretas em relação ao número total de amostras classificadas, em outras palavras, tem a capacidade em classificar corretamente as amostras considerando tanto os verdadeiros positivos quanto os verdadeiros negativos (Equação 15)<sup>144,147</sup>.

$$\text{Sensibilidade (\%)} = \left( \frac{TP}{TP+FN} \right) \times 100 \quad (13)$$

$$\text{Especificidade (\%)} = \left( \frac{TN}{TN+FP} \right) \times 100 \quad (14)$$

$$\text{Acurácia (\%)} = \left( \frac{TP+TN}{TP+FN+TN+FP} \right) \times 100 \quad (15)$$

onde TP (verdadeiro positivo) é igual ao número de amostras da classe de interesse (A) classificadas como sua verdadeira classe de interesse (A), TN (verdadeiro negativo) corresponde as amostras da classe de não interesse (B) classificadas como sua verdadeira classe de não interesse (B), FP (falso positivo) é o número de amostras da classe de não interesse (B) classificadas como da classe de interesse (A), e FN (falso negativo) é o número de amostras da classe de interesse (A) classificadas como da classe de não interesse (B)<sup>144,147</sup>. Essa forma de avaliação da eficiência do modelo considerando os quatro resultados possíveis é conhecida como coeficiente de correlação de Matthews<sup>148</sup>.

## **4. RESULTADOS E DISCUSSÃO**

### **4.1 CLASSIFICAÇÃO DE SUCOS DE UVA COM RESPEITO AO SISTEMA DE PRODUÇÃO AGRÍCOLA POR MEIO DE ESPECTROSCOPIA VISÍVEL ASSISTIDA POR QUIMIOMETRIA**

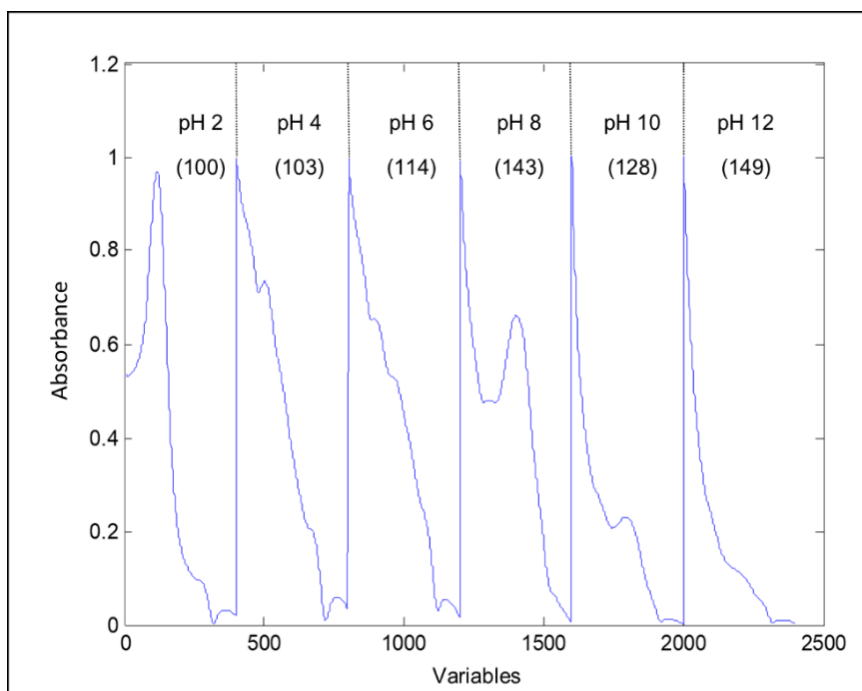
Esta seção tem como objetivo apresentar o estudo envolvendo a classificação de sucos de uva integral com respeito ao sistema de produção agrícola por meio de espectroscopia na região do visível assistida por Quimiometria (O artigo publicado encontra-se disponível em: <https://doi.org/10.1016/j.jfca.2022.104793>).

#### **4.1.1 Espectros Visíveis em Diferentes pHs**

Os conjuntos de dados espectrais na região do visível referentes às amostras de suco de uva, para cada pH diferente investigado, foram, primeiramente, submetidos a pré-tratamento por normalização, a fim de colocar os dados na mesma escala e reduzir variações indesejadas. A suavização pelo filtro Savitzky-Golay (suavização regular, polinômio de segunda ordem e tamanho da janela de 15 pontos) também foi realizada, a fim de melhorar a qualidade do sinal removendo ruídos instrumentais aleatórios, sem perder informações espectrais importantes<sup>149</sup>. Além dos dois procedimentos mencionados no tratamento dos dados, que ocasionaram melhora nos espectros visíveis, foi aplicado a primeira e segunda derivadas no conjunto de dados para melhorar a correção da linha de base, porém sem efeito favorável significativo.

Em seguida, após o pré-tratamento, os dados visíveis foram submetidos à modulação do pH (concatenação das matrizes de cada pH) de acordo com duas estratégias: (i) utilizando todas as soluções tampão (2, 4, 6, 8, 10 e 12), resultando em uma matriz de dados de  $46 \times 2400$ ; e (ii) utilizando apenas soluções tampão com pH alcalino (8, 10 e 12), resultando em uma matriz de dados de tamanho  $46 \times 1200$ . A escolha pelo segundo planejamento, de construir uma matriz de dados utilizando apenas soluções tampão com pHs alcalinos, se deve aos resultados obtidos na primeira etapa (com todos os pHs na modulação) observados na Figura 27, em que mais variáveis foram selecionadas nos pHs mais alcalinos pelos modelos GA-LDA.

**Figura 27.** Perfis espectrais médios visíveis após modulação de pH para todos os pHs estudados (2, 4, 6, 8, 10 e 12).



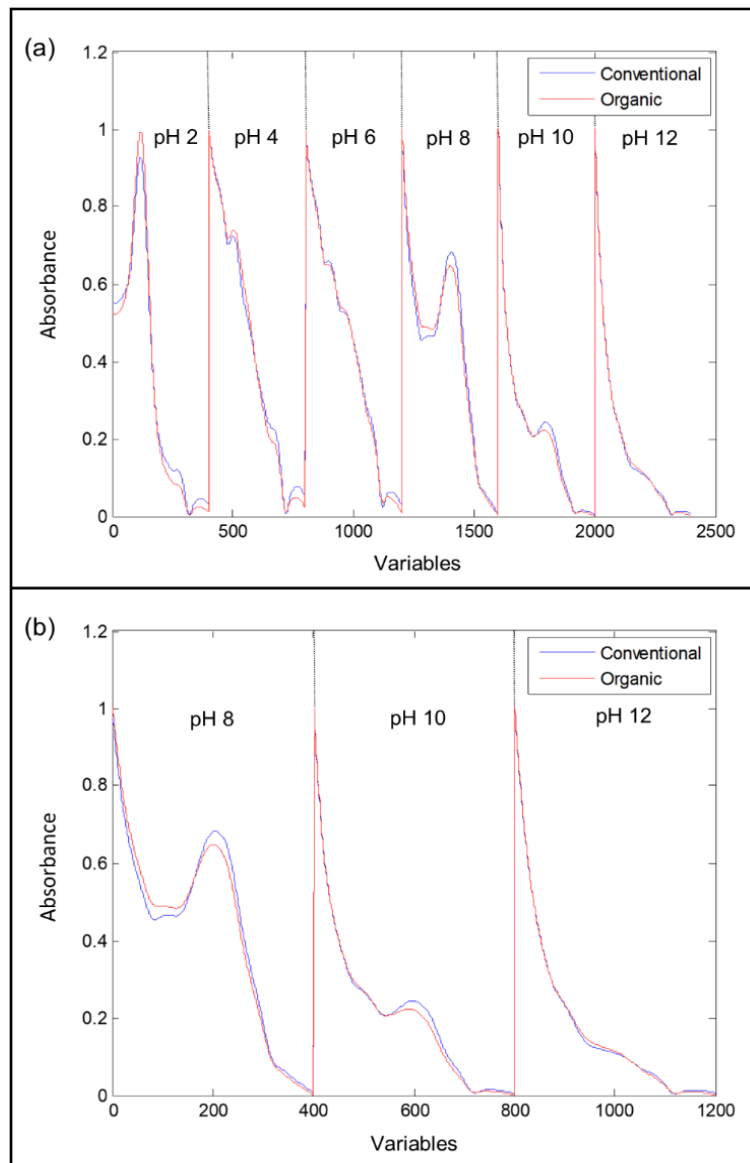
Nota: Em parênteses, o número total de variáveis selecionadas por todos os modelos GA-LDA.

Fonte: elaborada pelo autor, 2022.

Na primeira estratégia, foi somado o número total de variáveis selecionadas pelo GA para todos os modelos construídos (13 modelos, em triplicata, totalizando 39 modelos ao final), resultando em 100 (pH = 2), 103 (pH = 4), 114 (pH = 6), 143 (pH = 8), 128 (pH = 10) e 149 (pH = 12). É possível notar que para os pHs mais alcalinos, o GA selecionou mais variáveis. Considerando essa condição observada, segunda proposta foi adotada com a construção de novos modelos LDA (7 modelos, em quintuplicata, totalizando 35 modelos ao final), apenas para o conjunto de dados de pHs alcalinos e para os *inputs* que apresentaram melhor acurácia para o conjunto de teste (acima de 75%) e menores valores de amostras mal classificadas (até quatro amostras) atingidos na primeira estratégia.

A Figura 28 mostra as diferenças nos perfis espectrais visíveis após a modulação do pH para os conjuntos de amostras (convencional e orgânico) para todos os pHs analisados e apenas para os pHs alcalinos, respectivamente. É possível observar que para alguns pHs (como o 2, 4, 8 e 10) existem diferenças entre os espectros médios das classes estudadas, o que foi um indicativo de que a construção de modelos de classificação supervisionada pode ser bem sucedida.

**Figura 28.** Perfis espectrais médios visíveis após a modulação do pH das classes convencional e orgânica para (a) todos os pHs (2, 4, 6, 8, 10 e 12) e (b) pHs alcalinos (8, 10 e 12).



Fonte: elaborada pelo autor, 2022.

#### 4.1.2 Modelos LDA e PLS-DA

Os modelos GA-LDA foram construídos para classificação e seleção das variáveis mais significativas. As configurações de entrada utilizadas para produzir os diferentes modelos foram as seguintes: população de 50, 100 e 200; número de gerações de 100, 200 e 500; taxa de mutação de 5%, 10% e 15%; máximo de variáveis a serem selecionadas entre 15, 25 e 40; e taxa de cruzamento de 40%, 60% e 80%. Com os parâmetros

escolhidos, foi realizada análise para cada combinação, inicialmente utilizando dados com todos os pHs estudados.

Os resultados expressos na Tabela 8 mostram que o melhor modelo GA-LDA para todos os pHs apresentou taxas de 100% para sensibilidade, seletividade e acurácia do conjunto de treinamento para as classes convencional e orgânica. A partir do conjunto de testes, obtiveram-se taxas de 83,3% de sensibilidade e 88,9% de seletividade para a classe convencional e taxas de 88,9% de sensibilidade e 83,3% de seletividade para a classe orgânica. A taxa de acurácia para ambas as classes convencional e orgânica foi de 86,1%.

O melhor modelo GA-LDA para pHs alcalinos apresentou resultado ligeiramente superior com índices de 100% para sensibilidade, seletividade e acurácia do conjunto de treinamento para as classes convencional e orgânica. A partir do conjunto de testes foram obtidos índices de 83,3% de sensibilidade e 100% de seletividade para a classe convencional e índices de 100% de sensibilidade e 83,3% de seletividade para a classe orgânica. O índice de acerto para ambas as classes convencional e orgânica foi de 91,7%.

Comparativamente, o modelo PLS-DA para os dados com todos os pHs apresentou como resultados as taxas de 83% de sensibilidade e 95% de seletividade para a classe convencional e taxas de 95% de sensibilidade e 83% de seletividade para a classe orgânica do conjunto de treinamento. A taxa de acurácia foi de 90% para ambas as classes convencional e orgânica. A partir do conjunto de testes, obtiveram-se taxas de 83% de sensibilidade e 56% de seletividade para a classe convencional e taxas de 56% de sensibilidade e 83% de seletividade para a classe orgânica. A taxa de acurácia foi de 67% para ambas as classes convencional e orgânica.

Por fim, o resultado do modelo PLS-DA para pHs alcalinos apresentou índices de 100% para sensibilidade, seletividade e acurácia do conjunto de treinamento para as classes convencional e orgânica. A partir do conjunto de testes, foram obtidos índices de 83% de sensibilidade e 67% de seletividade para a classe convencional e índices de 67% de sensibilidade e 83% de seletividade para a classe orgânica. O índice de acurácia para as ambas as classes convencional e orgânica foi de 73%. Os resultados, em geral, revelaram um desempenho inferior nas figuras de mérito do modelo PLS-DA em relação aos modelos GA-LDA, tanto para todos os pHs quanto para os pHs alcalinos estudados.



**Tabela 8.** Figuras de mérito dos melhores modelos de classificação para dados de espectroscopia na região do visível.

Modelo GA-LDA para todos os pHs (2, 4, 6, 8, 10 e 12)					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	5	1
	Orgânico	0	19	1	8
Sensibilidade (%)		100	100	83,3	88,9
Especificidade (%)		100	100	88,9	88,3
Acurácia (%)		100		86,1	
Modelo GA-LDA para pHs alcalinos (8, 10 e 12)					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	5	1
	Orgânico	0	19	0	9
Sensibilidade (%)		100	100	83,3	100
Especificidade (%)		100	100	100	83,3
Acurácia (%)		100		91,7	
Modelo PLS-DA para todos os pHs (2, 4, 6, 8, 10 e 12) com 5 variáveis latentes					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	10	2	5	1
	Orgânico	1	18	4	5
Sensibilidade (%)		83	95	83	56
Especificidade (%)		95	83	56	83
Acurácia (%)		90		67	
Modelo PLS-DA para pHs alcalinos (8, 10 e 12) com 7 variáveis latentes					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	5	1
	Orgânico	0	19	3	6
Sensibilidade (%)		100	100	83	67
Especificidade (%)		100	100	67	83
Acurácia (%)		100		73	

Fonte: elaborada pelo autor, 2022.

O modelo com melhor resultado para todas as soluções tampão empregou o seguinte *input*: 100 indivíduos, 200 gerações, no máximo 25 variáveis selecionadas, taxa de mutação de 5% e cruzamento de 60%. O melhor resultado obtido com soluções tampão com pHs alcalinos utilizou 100 indivíduos, 500 gerações, no máximo 15 variáveis selecionadas, taxa de mutação de 10% e cruzamento de 80%. O modelo GA-LDA com dados selecionados para pHs alcalinos obteve um desempenho melhor do que o melhor modelo para todos os pHs; entretanto, ambos os modelos obtiveram alta acurácia para validação e predição.

Um modelo PLS-DA foi construído com todo o conjunto de dados para comparar o desempenho com os modelos GA-LDA. Os resultados revelaram uma menor acurácia com perda de sensibilidade para a classe orgânica ao avaliar o conjunto de dados total. Para os dados de pH alcalino, o modelo PLS-DA apresentou melhora na taxa de classificação e nos demais indicadores, porém ainda abaixo das soluções obtidas dos modelos GA-LDA. Menor desempenho para os modelos PLS-DA também foi observado para os dados de cada pH individual, conforme a Tabela 9.

**Tabela 9.** Figuras de mérito dos modelos PLS-DA para cada pH individual para dados de espectroscopia na região do visível.

Modelo PLS-DA para pH 2 com 7 variáveis latentes					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	9	3	3	3
	Orgânico	2	17	3	6
Sensibilidade (%)		75	89	50	67
Especificidade (%)		89	75	67	50
Acurácia (%)		84		60	
Modelo PLS-DA para pH 4 com 7 variáveis latentes					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	3	3
	Orgânico	2	17	3	6
Sensibilidade (%)		100	89	50	67
Especificidade (%)		89	100	67	50
Acurácia (%)		94		60	
Modelo PLS-DA para pH 6 com 6 variáveis latentes					

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	10	2	5	1
	Orgânico	3	16	3	6
Sensibilidade (%)		83	84	83	67
Especificidade (%)		84	83	67	83
Acurácia (%)		84		73	

Modelo PLS-DA para pH 8 com 8 variáveis latentes

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	9	3	6	0
	Orgânico	3	16	2	7
Sensibilidade (%)		75	84	100	78
Especificidade (%)		84	75	78	100
Acurácia (%)		81		87	

Modelo PLS-DA para pH 10 com 4 variáveis latentes

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	9	3	6	0
	Orgânico	1	18	1	8
Sensibilidade (%)		75	95	100	89
Especificidade (%)		95	75	89	100
Acurácia (%)		87		93	

Modelo PLS-DA para pH 12 com 6 variáveis latentes

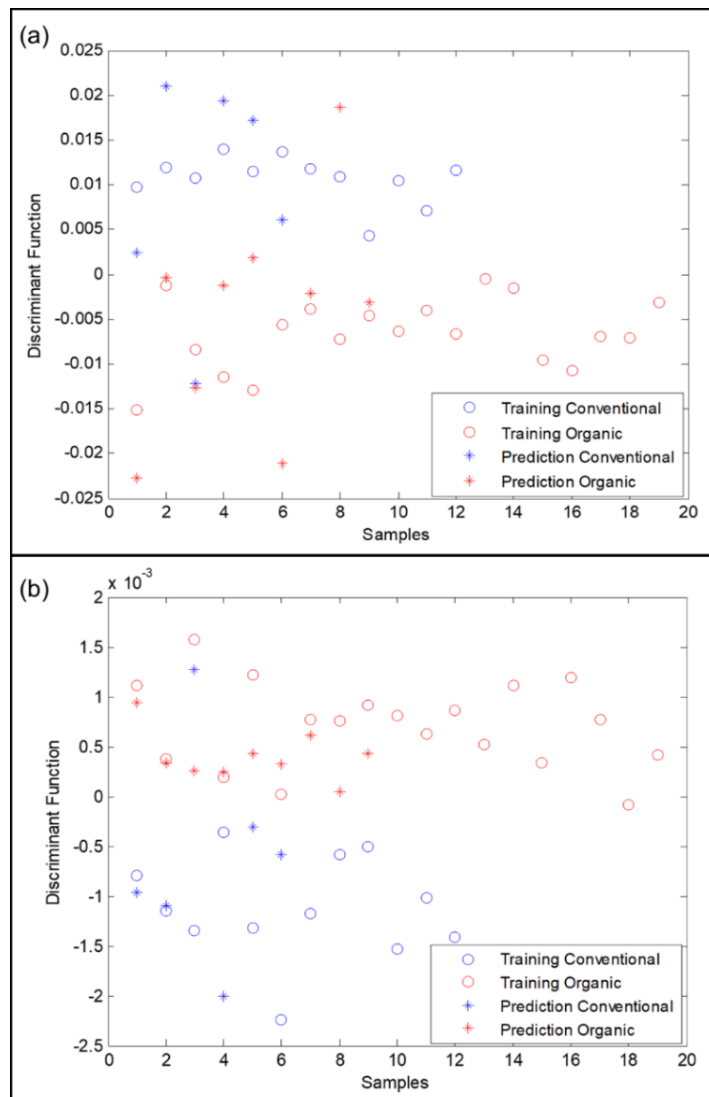
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	11	1	5	1
	Orgânico	2	17	2	7
Sensibilidade (%)		92	89	83	78
Especificidade (%)		89	92	78	83
Acurácia (%)		90		80	

Fonte: elaborada pelo autor, 2022.

Gráficos de função discriminante para os dois estudos estabelecidos (todos os pHs e apenas pHs alcalinos) são reproduzidos na Figura 29. As amostras de treinamento para os

sucos produzidos no sistema convencional (círculos azuis) e teste (estrelas azuis), quando se utilizou a modulação com todos os pHs, apresentaram valores positivos para a função discriminante. Em contraste, as amostras de treinamento para os sucos orgânicos (círculos vermelhos) e teste (estrelas vermelhas) apresentam valores negativos para a função discriminante. Os resultados revelaram que duas amostras de teste foram erroneamente classificadas como convencional e orgânica. Para o conjunto com apenas as soluções tampão alcalinas, o perfil do gráfico de função discriminante para as amostras de treinamento e teste é invertido. Os resultados mostraram uma amostra de teste da classe convencional agrupada como orgânica, um falso positivo com valor acima da função discriminante.

**Figura 29.** Gráficos de função discriminante das amostras de treinamento e teste dos melhores modelos por GA-LDA para (a) todos os pHs (2, 4, 6, 8, 10 e 12) e (b) pHs alcalinos (8, 10 e 12).



Fonte: elaborada pelo autor, 2022.

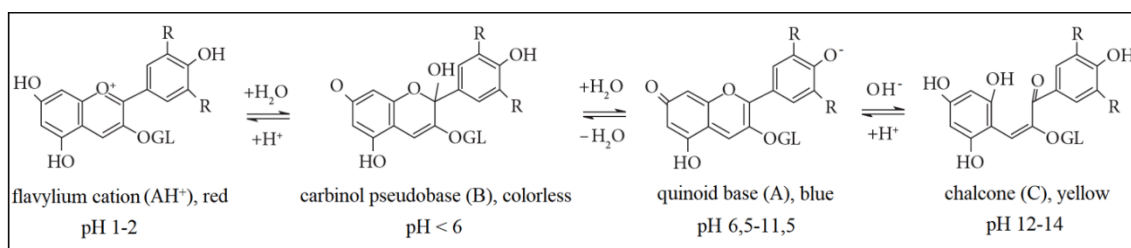
### 4.1.3 Seleção de Variáveis

O suco de uva possui uma composição química que absorve radiação eletromagnética em uma ampla faixa do espectro visível, indicando um perfil espectroscópico diferente de acordo com o valor de pH e exibindo bandas em diferentes regiões e com intensidade variável. Os comprimentos de onda de absorção observados representam o perfil de compostos fenólicos presentes no suco de uva e a quantidade relativa desses compostos varia de acordo com o sistema agrícola aplicado.

Os compostos polifenólicos fazem parte do sistema de defesa da planta e são formados na fase de maturação da uva. O conteúdo é influenciado pelas condições de cultivo que afetam a absorção e disponibilidade de compostos fenólicos na estrutura da parede celular da planta, que se prevê serem maiores em plantas cultivadas pelo sistema orgânico. A produção de sucos orgânicos é definida pela ausência de agrotóxicos e um menor uso de fertilizantes, portanto, as plantas produzidas organicamente têm um período de maturação mais longo em relação às uvas convencionais devido à liberação mais lenta dos nutrientes<sup>150,151</sup>.

As bandas observadas no espectro são essencialmente características das antocianinas, um subgrupo de flavonóides que, dependendo do nível de acidez ou alcalinidade, assumem diferentes estruturas químicas, incluindo o cátion flavílio ( $AH^+$ ), pseudobase carbinol (B), base quinóide (A) e chalcona (C), conforme mostra a Figura 30<sup>152-154</sup>. As antocianinas são facilmente distinguíveis por sua forte absorbância na região visível do espectro em torno de 520 nm, incluindo vários compostos como cianidina 3,5-diglicosídeo, malvidina 3-O-glicosídeo, malvidina 3,5-diglicosídeo, pelargonidina 3,5-diglicosídeo e peonidina 3-O-glicosídeo<sup>13</sup>.

**Figura 30.** Alterações estruturais de uma antocianina genérica em meio aquoso em função da variação do pH.



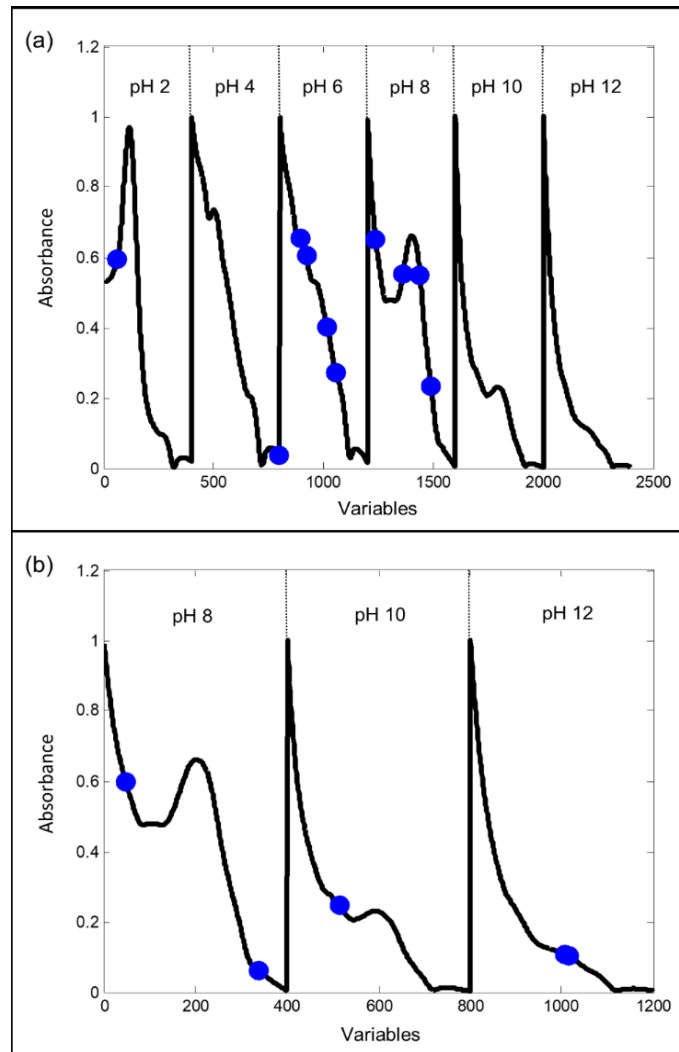
Fonte: adaptada de Levi, 2004<sup>153</sup>.

As antocianinas podem se associar ou formar copigmentos em reação com outros compostos orgânicos incolores, principalmente alcalóides, aminoácidos, flavonóides não antociânicos e ácidos fenólicos, quando presentes como espécies de flavílios<sup>155</sup>. A co-pigmentação geralmente resulta em aumento na absorbância e deslocamento positivo do pico máximo na absorbância visível<sup>156</sup>. As regiões espectrais visíveis selecionadas podem estar relacionadas a reações enzimáticas e não enzimáticas durante a desidratação da uva levando a pigmentos escurecidos. O escurecimento enzimático está associado ao teor de ácidos caftárico e coutárico, promovido por flavonóides, enquanto o escurecimento não enzimático pode surgir da oxidação e polimerização<sup>157,158</sup>. Os sinais em comprimentos de onda em torno de 420, 450 e 520 nm são particulares e referem-se ao índice de escurecimento, pigmentos de xantílio e cor do pigmento polimérico, respectivamente<sup>159</sup>.

As variáveis mais relevantes selecionadas pelo GA a partir dos melhores modelos LDA, correspondentes aos comprimentos de onda de absorção em diferentes pHs, são identificadas como pontos azuis na Figura 31. Para o conjunto de dados envolvendo todos os pHs, as variáveis obtidas do melhor modelo foram: 462 nm (pH = 2); 795 nm (pH = 4); 493, 523, 617 e 655 nm (pH = 6); e 438, 561, 640 e 690 nm (pH = 8). Da mesma forma, para os dados referentes aos pHs básicos, os comprimentos de onda do melhor modelo foram: 448 e 739 nm (pH = 8); 515 nm (pH = 10); e 609 e 616 nm (pH = 12).

Em geral, em meio extremamente ácido (pH entre 1 e 2), há predominância de antocianinas na forma catiônica flavílio (2-fenilbenzopirílio) (cor vermelha), que tem absorção máxima em torno de 510 nm, variando a valores de absorbância mais baixos<sup>160</sup>. O sinal selecionado em 462 nm em pH = 2 pode se referir ao cátion flavílio ou pigmentos derivados de xantílio. Para um meio com pH entre 2 e 6, há um equilíbrio ácido-base de protonação da estrutura do cátion flavílio e uma formação conhecida como carbinol (incolor) é observada<sup>161</sup>. O sinal em 795 nm em pH = 4 pode estar relacionada a reações de co-pigmentação. Com a diminuição da acidez, em soluções tampão fracamente ácidas e neutras, o espectro não mostra mais a banda de absorção da estrutura do flavílio<sup>153</sup>. Sinais em 493, 523, 617 e 655 nm podem estar associados a inúmeros pigmentos poliméricos derivados de antocianinas, como piranoantocianinas dos tipos flavanil, pinotina e vitisina, ou mono- ou diglicosídeos e moléculas acetiladas e com reações de co-pigmentação<sup>162,163</sup>.

**Figura 31.** Seleção das variáveis mais relevantes dos melhores modelos por GA-LDA para (a) todos os pHs (2, 4, 6, 8, 10 e 12) e (b) pHs alcalinos (8, 10 e 12).



Fonte: elaborada pelo autor, 2022.

Em meio levemente neutro, ocorre efeito batocrômico em 540 nm devido à formação da anidrobases quinoidal e a presença de outro sinal na região de 440 nm, definida como chalcona ionizada, com sinais observados em 438 e 448 nm em pH = 8. O sinal em 561 nm pode estar relacionada à base quinoidal e os sinais em 640, 690 e 739 nm podem se referir a formas quinonas de antocianinas livres<sup>164</sup> ou reações de co-pigmentação<sup>156</sup>. Em meio extremamente alcalino, a absorção máxima na faixa de 600 nm é comumente observada, correspondendo à base quinoidal ionizada, deslocada em relação aos extratos ácidos, mas com menor intensidade. O sinal em 465 nm, balanceada entre formas ionizadas de chalconas, mostra uma coloração amarelada da solução<sup>153,160</sup>. Os sinais em 609 e 616 nm podem ser associados a esta estrutura de base quinoidal ionizada e o sinal em 515 nm pode ser correlacionado com a forma de chalcona ionizada ou a pigmentos poliméricos.

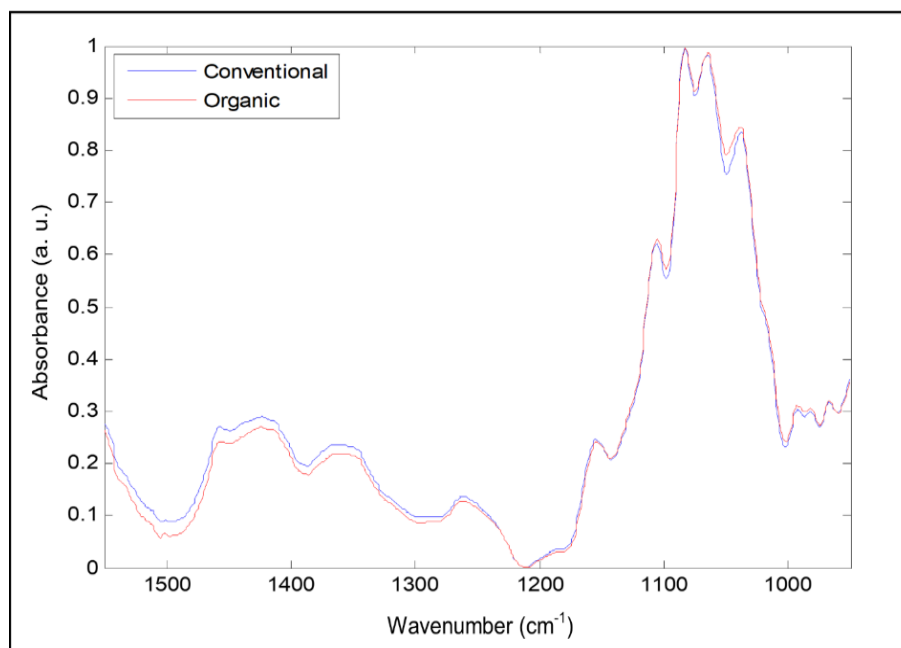
## 4.2 METODOLOGIA ANALÍTICA VERDE PARA CLASSIFICAÇÃO DE SUCOS DE UVA USANDO ESPECTROSCOPIA FTIR COMBINADA COM QUIMIOMETRIA

Esta seção teve propósito de avaliar a utilização de dados espectroscópicos na região do infravermelho médio para classificação de sucos de uva integral seguindo os princípios estabelecidos da Química Analítica Verde (QAV) baseado no conceito de sustentabilidade com mínimo impacto ambiental (O artigo publicado encontra-se disponível em: <https://doi.org/10.1016/j.talo.2022.100168>).

### 4.2.1 Espectros no Infravermelho Médio

Os conjuntos de dados espectrais na região do infravermelho médio (FTIR) foi, inicialmente, subordinado a pré-tratamento pelo algoritmo Savitzky-Golay (polinômio de 2ª ordem e tamanho da janela de 13 pontos) para suavização para remover ruído instrumental aleatório, após normalização para colocar os dados na mesma escala, e realizado a média da triplicata dos dados. No conjunto de dados, também, foi aplicado os tratamentos da primeira e segunda derivadas, no entanto sem melhora espectral. A Figura 32 mostra os perfis espectrais no infravermelho médio dos valores médios para as duas classes de amostra (convencional e orgânica).

**Figura 32.** Perfis espectrais FTIR dos valores médios das classes convencional e orgânica.



Fonte: elaborada pelo autor, 2022.



#### 4.2.2 Modelos LDA e PLS-DA

Os modelos LDA para discriminação e seleção das variáveis mais significativas foram realizados pelos algoritmos GA, SW, SPA e ACO. As configurações utilizadas para construir os diferentes modelos foram as seguintes: para GA, população de 50, 100 e 200; número de gerações de 100, 200 e 500; taxa de mutação de 5%, 10% e 15%; número máximo de variáveis selecionadas de 15, 25 e 40; e taxa de cruzamento de 40%, 60% e 80%; para SW, correlação máxima de 0,2 a 0,9 com variação de 0,5; para SPA, tamanho de cadeia de variáveis de 1 a 15; e para ACO, o número de formigas foi 50, 100 e 200; número de colônias de 100, 200 e 500; taxa de formigas cegas de 0,25%, 0,35%, 0,50% e 0,70%; número máximo de variáveis selecionadas de 15 e 25; e taxa de evaporação de feromônio de 40% e 65%.

A análise foi realizada para cada combinação escolhida dos parâmetros em triplicata nos métodos GA e ACO, que são algoritmos meta-heurísticos, ou seja, utilizam um processo aleatório para encontrar a melhor solução possível (otimização estocástica). A técnica PLS-DA também foi aplicada para avaliar seu desempenho com os modelos LDA. Os resultados mostraram que todas as amostras foram categorizadas com excelente taxa de classificação correta (TCC) com 100% nos conjuntos de treinamento e teste ao usar a ferramenta LDA com os diferentes algoritmos de seleção de variáveis para classificar sucos de uva integral de acordo com o sistema de produção agrícola. O modelo PLS-DA também mostrou o resultado de 100% de TCC para os conjuntos de treinamento e teste.

A técnica LDA proposta mostrou confiabilidade na obtenção de soluções análogas ao resultado obtido pelo conhecido modelo PLS-DA. Isso prova que o conjunto de dados é robusto devido à sua capacidade de ser aplicado em diferentes técnicas de classificação supervisionadas por análise discriminante, apresentando como resultados modelos com soluções satisfatórias. As figuras de mérito e matrizes de confusão com desempenhos para cada metodologia e algoritmos são apresentadas na Tabela 10.

**Tabela 10.** Figuras de mérito dos melhores modelos de classificação para dados de espectroscopia na região do infravermelho médio.

Modelo GA-LDA com 6 variáveis latentes				
Classe Atribuída	Treinamento		Teste	
	Convencional	Orgânico	Convencional	Orgânico

Matriz de Confusão	Convencional	12	0	6	0
	Orgânico	0	18	0	9
Sensibilidade (%)		100	100	100	100
Especificidade (%)		100	100	100	100
Acurácia (%)		100		100	

Modelo SW-LDA com 5 variáveis latentes

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	6	0
	Orgânico	0	18	0	9
Sensibilidade (%)		100	100	100	100
Especificidade (%)		100	100	100	100
Acurácia (%)		100		100	

Modelo SPA-LDA com 3 variáveis latentes

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	6	0
	Orgânico	0	18	0	9
Sensibilidade (%)		100	100	100	100
Especificidade (%)		100	100	100	100
Acurácia (%)		100		100	

Modelo ACO-LDA com 7 variáveis latentes

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	6	0
	Orgânico	0	18	0	9
Sensibilidade (%)		100	100	100	100
Especificidade (%)		100	100	100	100
Acurácia (%)		100		100	

Modelo PLS-DA com 2 variáveis latentes

Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
	Convencional	12	0	6	0

Matriz de Confusão	Orgânico	0	18	0	9
Sensibilidade (%)		100	100	100	100
Especificidade (%)		100	100	100	100
Acurácia (%)		100		100	

Fonte: elaborada pelo autor, 2022.

O melhor modelo GA-LDA empregou os seguintes parâmetros: 100 populações, 500 gerações, máximo de 15 variáveis e taxas de mutação e crossover de 60% e 10%, respectivamente. O método SW-LDA apresentou o critério de 0,5 como correlação máxima, e o modelo SPA-LDA utilizou um tamanho de cadeia variável de 1 (mínimo de variáveis) a 15 (máximo de variáveis) para as etapas de projeção. Para o melhor modelo ACO-LDA, os critérios utilizados foram 100 para o número de formigas, 500 para o número de colônias, máximo de 15 variáveis selecionadas, taxa de formigas cegas de 0,5% e taxa de evaporação de feromônio de 40%.

Ao pesquisar na literatura, há poucas pesquisas a respeito da classificação dos sucos de uva integral com relação ao sistema agrícola de cultivo a fim de verificar a autenticidade do suco orgânico. Dentre esses estudos encontrados, temos os seguintes resultados: taxas de acurácia de 100% do conjunto de treinamento e 92% do conjunto de teste por meio do uso de espectroscopia na região do visível com técnicas LDA e PLS-DA<sup>165</sup>; 89% de precisão usando validação cruzada leave-one-out usando ICP-MS com máquinas de vetores de suporte, redes neurais e árvores de decisão<sup>166</sup>; taxas de seletividade de 82% para classe convencional e 95% para classe orgânica empregando ICP-MS e SIMCA<sup>167</sup>; índices de precisão de 78% para calibração e 75% para validação utilizando dados cromatográficos e espectrofotométricos de composição química com PLS-DA<sup>168</sup>; e taxas de 81% de seletividade para classe convencional e 83% para classe orgânica usando técnicas PTR-MS e PLS-DA<sup>169</sup>.

As soluções obtidas desses trabalhos confirmaram desempenhos inferiores das figuras de mérito aplicadas sobre os modelos em comparação com o que foi demonstrado pelo presente estudo utilizando dados espectrais FTIR com modelos LDA e PLS-DA. Além disso, todas essas pesquisas mencionadas apresentaram mais desvantagens quando relacionadas aos princípios da Química Analítica Verde (QAV), em que requereram utilização de solventes com formação de resíduos. Os objetivos da QAV enfatizam a necessidade de reduzir ou eliminar o uso de solventes, buscar solventes mais seguros e

menos tóxicos, geração mínima de resíduos e baixo consumo de energia<sup>170,171</sup>. O conjunto dos 12 conceitos estabelecidos pela QAV são os seguintes<sup>172</sup>:

1. **Prevenção de resíduos:** é melhor prevenir resíduos do que tratar ou limpar os resíduos depois de criados.

2. **Eficiência atômica:** métodos sintéticos devem ser desenvolvidos para maximizar a incorporação de todos os materiais usados no processo no produto final.

3. **Sínteses químicas menos perigosas:** sempre que praticável, a síntese do produto químico deve usar e gerar substâncias que possuem pouca ou nenhuma toxicidade para a saúde humana e o meio ambiente.

4. **Desenvolvimento de produtos químicos mais seguros:** produtos químicos devem ser projetados para preservar a eficácia da função enquanto reduzem a toxicidade.

5. **Solventes e auxiliares mais seguros:** uso de substâncias auxiliares (solventes, agentes de separação) deve ser desnecessário sempre que possível e inócuo quando usado.

6. **Busca pela eficiência energética:** requisitos de energia devem ser reconhecidos por seus impactos ambientais e econômicos e devem ser minimizados. Os métodos sintéticos devem ser conduzidos à temperatura e pressão ambiente.

7. **Uso de matérias-primas renováveis:** sempre que técnica e economicamente viável, a utilização de matérias-primas renováveis deve ser escolhida em detrimento de fontes não-renováveis.

8. **Evitar a formação de derivados:** derivatização desnecessária (uso de grupos de bloqueio, proteção/desproteção, modificação temporária de processos físicos/químicos) deve ser minimizada ou evitada se possível, pois tais etapas requerem reagentes adicionais e podem gerar resíduos.

9. **Catálise:** reagentes catalíticos (tão seletivos quanto possível) são superiores aos reagentes estequiométricos.

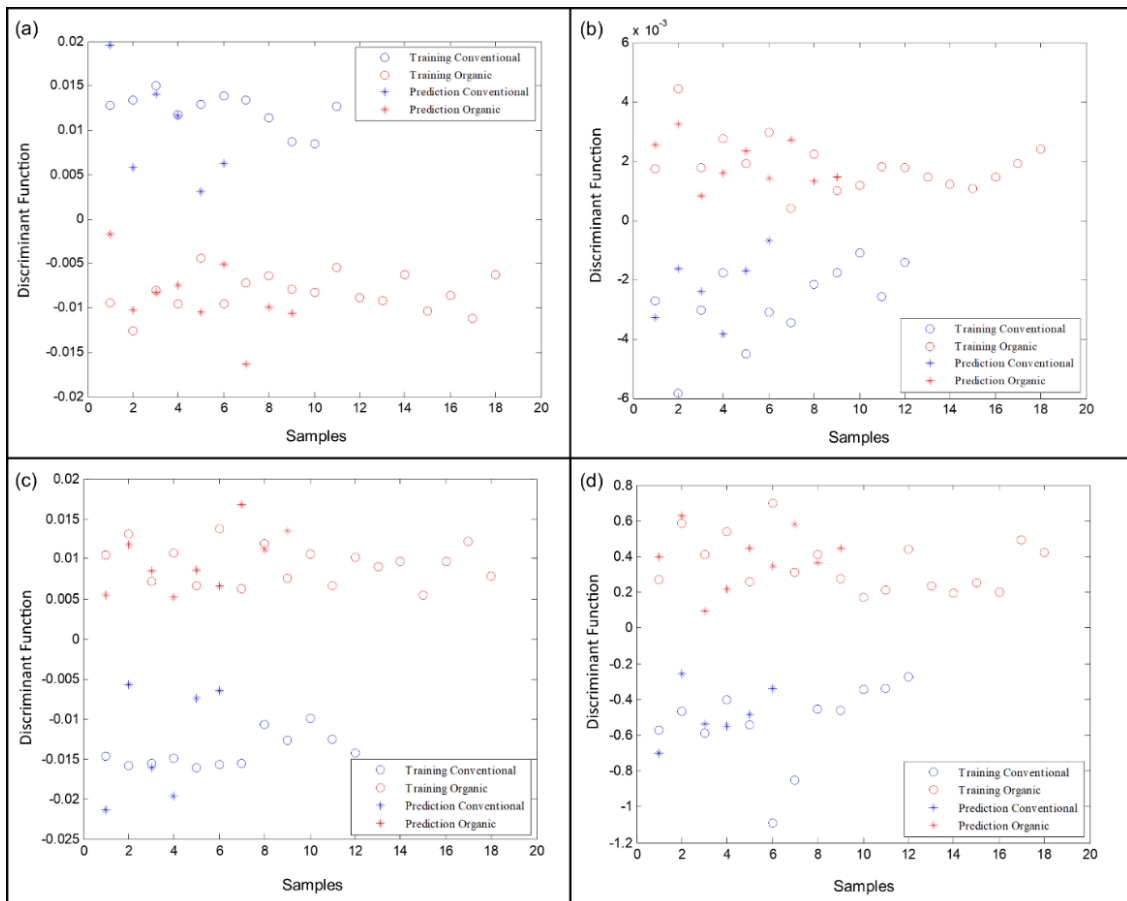
10. **Produtos degradáveis:** produtos químicos devem ser projetados de modo que, ao final de sua função, se decomponham em produtos de degradação inócuos e não persistam no meio ambiente.

11. **Análise em tempo real para prevenção da poluição:** metodologias analíticas precisam ser mais desenvolvidas para permitir o monitoramento e controle em tempo real em processo antes da formação de substâncias perigosas.

12. **Química inerentemente mais segura para prevenção de acidentes:** escolha das substâncias, bem como sua utilização em um processo químico, devem procurar a minimização do risco de acidentes, como vazamentos, incêndios e explosões.

Os gráficos das funções discriminantes para as duas classes de modelos escolhidas para cada algoritmo são apresentados na Figura 33. As amostras de treinamento convencional (círculos azuis) e teste (estrelas azuis) do modelo GA-LDA exibiram uma função discriminante com valores acima de zero. Em contraste, as amostras orgânicas de treinamento (círculos vermelhos) e teste (estrelas vermelhas) apresentaram uma função discriminante abaixo de zero. Para os modelos SW-, ACO- e SPA-LDA, o perfil dos gráficos de função discriminante para as amostras de treinamento e teste são invertidos. Os resultados mostraram que para todos os algoritmos avaliados, as amostras de treinamento e teste foram classificadas com uma taxa de acerto de 100%.

**Figura 33.** Gráficos de função discriminante das amostras de treinamento e teste dos melhores modelos por (a) GA-LDA, (b) SW-LDA, (c) ACO-LDA e (d) SPA-LDA.



Fonte: elaborada pelo autor, 2022.

### 4.2.3 Seleção de Variáveis

A clara separação das duas classes (convencional e orgânica) dos sucos de uva integral pode ser explicada pelos discriminantes de Fisher mostrados na Figura 34. Os discriminantes de Fisher referem-se aos números de onda ( $\text{cm}^{-1}$ ) mais relevantes selecionados (linhas em azul) pelas abordagens de seleção de variáveis acopladas ao modelo LDA, e suas intensidades mostram o quanto elas contribuem para permitir a discriminação de cada classe. Para o modelo GA-LDA (Figura 34 (a)), os números de onda que apresentam a função discriminante acima de zero estão relacionados à classe convencional, enquanto que os abaixo de zero representam a classe orgânica. Por outro lado, os modelos SW- (Figura 34 (c)), ACO- (Figura 34 (e)) e SPA-LDA (Figura 34 (g)) indicam que os números de onda com função discriminante maior que zero referem-se à classe orgânica e menor que zero correspondem à classe convencional.

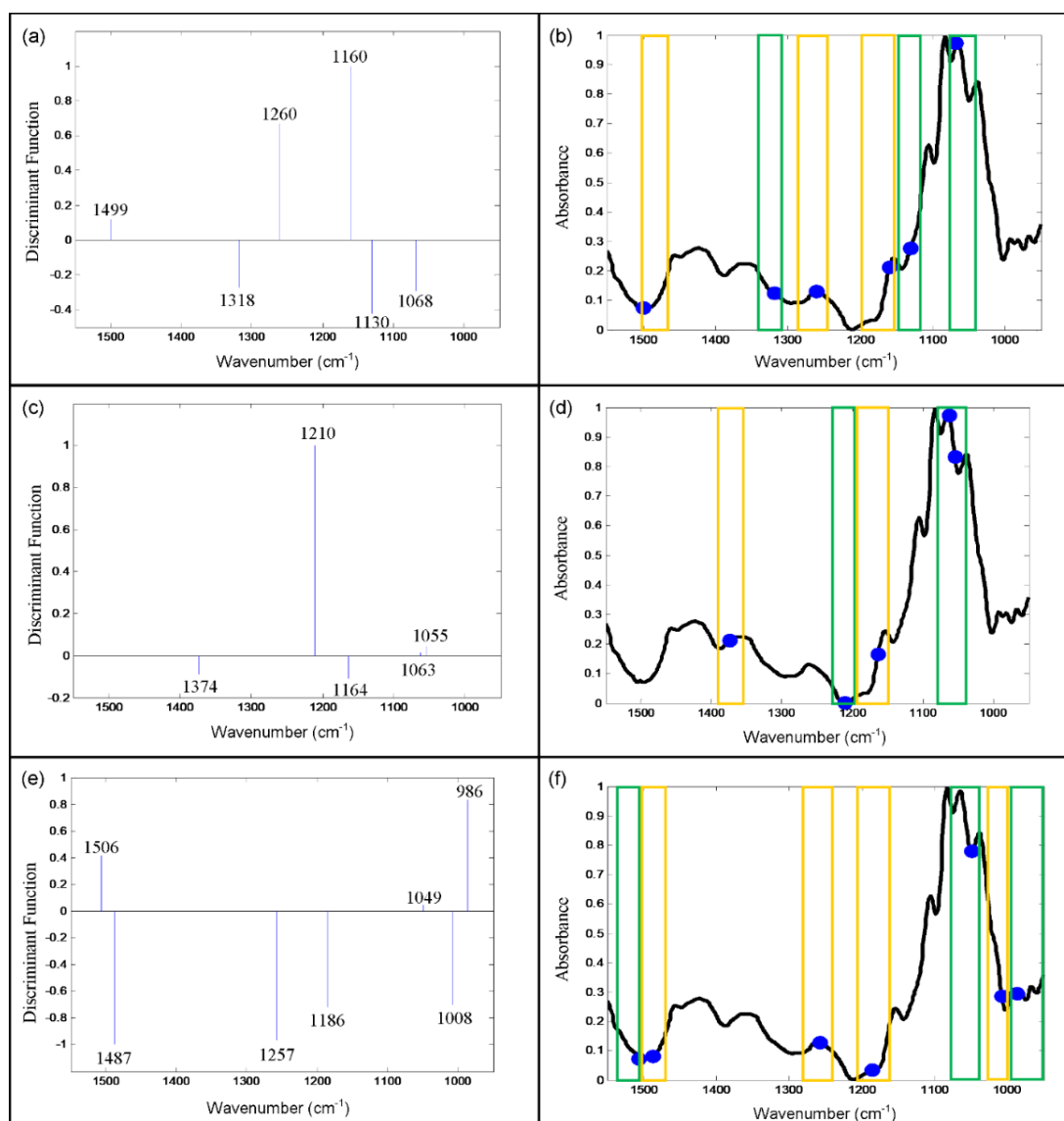
O perfil espectral com os números de onda selecionados por cada algoritmo GA (Figura 34 (b)), SW (Figura 34 (d)), ACO (Figura 34 (f)) e SPA (Figura 34 (h)) dos modelos LDA também pode ser visto na Figura 34. A seleção dos números de onda (grupos funcionais) mais relevantes pelos modelos permitiu distinguir regiões do espectro que servem como marcadores para classificação das amostras nas classes convencional e orgânica. Desta forma, podem ser identificadas as regiões do espectro que caracterizam as classes, onde as regiões de cor laranja correspondem à classe convencional e as regiões de cor verde representam a classe orgânica. Os pontos em azul no espectro representam os números de onda selecionados pelos algoritmos.

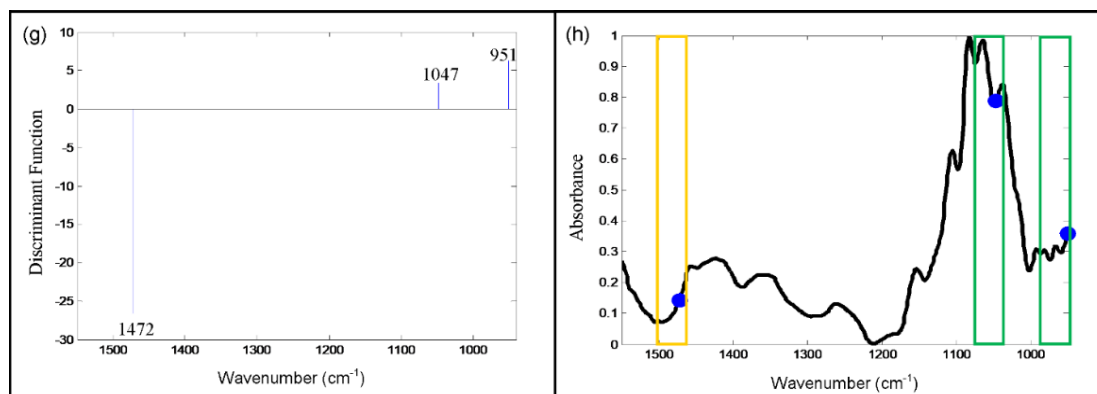
A identificação da absorção das principais bandas do espectro FTIR e suas respectivas atribuições são baseadas em dados da literatura. A região entre  $1550 \text{ cm}^{-1}$  e  $950 \text{ cm}^{-1}$  contém bandas típicas para compostos fenólicos, açúcares, ácidos orgânicos e substâncias nitrogenadas, como a vibração das ligações C–O, C–C, C–N e C–H<sup>173,174</sup>. Bandas características para compostos fenólicos são observadas entre  $1700 \text{ cm}^{-1}$  e  $1000 \text{ cm}^{-1}$  como o estiramento C=C do anel aromático ( $1510\text{-}1450 \text{ cm}^{-1}$ ), a deformação do fenol OH ( $1410\text{-}1310 \text{ cm}^{-1}$ ), a deformação no plano aromático C–H ( $1225\text{-}950 \text{ cm}^{-1}$ ) e o estiramento C–O do fenol ( $1200 \text{ cm}^{-1}$ )<sup>175</sup>.

Bandas de baixa intensidade entre  $1500 \text{ cm}^{-1}$  e  $1200 \text{ cm}^{-1}$  estão associadas a deformações de CH<sub>2</sub>, C–C–H e H–C–O de açúcares e ácidos orgânicos. As vibrações encontradas na região entre  $1100 \text{ cm}^{-1}$  e  $900 \text{ cm}^{-1}$  (impressão digital), com absorção máxima próxima a  $1050 \text{ cm}^{-1}$ , estão relacionadas ao estiramento das ligações C–C e C–O,

que correspondem à presença de açúcares (glicose e frutose) e ácidos orgânicos (principalmente ácido málico)<sup>175-178</sup>. Além disso, a região entre 1100  $\text{cm}^{-1}$  e 995  $\text{cm}^{-1}$  pode ser atribuída à deformação OH e estiramento C–O em compostos fenólicos, à deformação C–H dos grupos  $\text{CH}_2$  e  $\text{CH}_3$  e ao estiramento assimétrico  $\text{COO}^-$  de ésteres<sup>175,179</sup>. Portanto, a identificação dos grupos funcionais dos discriminantes de Fisher mais intensos que caracterizam os compostos orgânicos, principalmente as substâncias fenólicas, presentes nos sucos de uva é relevante para permitir a classificação das amostras nas classes convencional e orgânica.

**Figura 34.** Discriminantes de Fisher com seleção dos números de onda ( $\text{cm}^{-1}$ ) mais relevantes e o perfil espectral com as variáveis selecionadas para as classes convencional (regiões laranjas) e orgânica (regiões verdes) dos melhores modelos por GA-LDA (a) e (b), SW-LDA (c) e (d), ACO-LDA (e) e (f), e SPA-LDA (g) e (h), respectivamente.





Fonte: elaborada pelo autor, 2022.

Os sucos de uva são alimentos complexos devido ao grande número de compostos orgânicos presentes. Os compostos fenólicos são formados na fase de maturação da uva e atuam no sistema de defesa da videira, cujo teor é resultado do processo de produção agrícola aplicado. As uvas produzidas pelo sistema de agricultura orgânica são caracterizadas pela ausência do uso de agrotóxicos sintéticos para controle de pragas e de doenças na planta e, por esse motivo, apresentam um período de maturação mais longo devido à liberação mais lenta dos nutrientes em relação às uvas plantadas pela técnica agrícola convencional. As condições de cultivo afetam a absorção e disponibilidade de compostos fenólicos na estrutura da parede celular, que se prevê serem maiores em plantas cultivadas pelo sistema de agricultura orgânica<sup>150,180,181</sup>.

#### 4.3 DISCRIMINAÇÃO DE SUCO DE UVA INTEGRAL USANDO DADOS DE ESPECTROSCOPIA DE FLUORESCÊNCIA COM ANÁLISE DISCRIMINANTE LINEAR ACOPLADA AOS ALGORITMOS GENÉTICO E OTIMIZAÇÃO POR COLÔNIA DE FORMIGAS

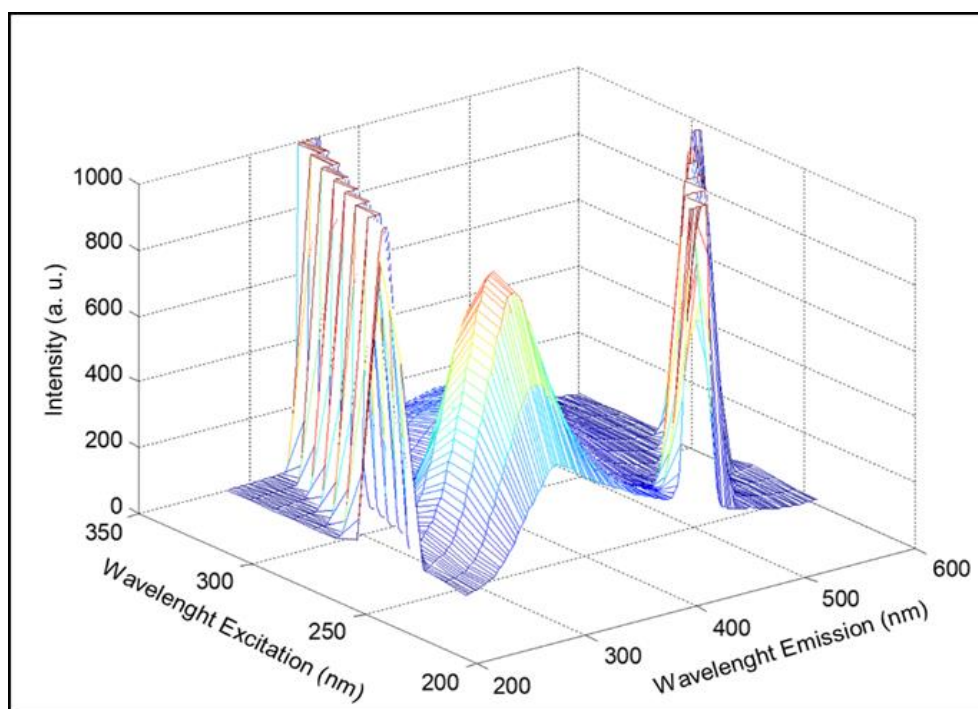
Esta seção teve como abordagem classificar sucos de uva usando a espectroscopia de fluorescência molecular combinada com LDA e seleção das variáveis mais importantes pelos algoritmos GA e ACO (O artigo publicado encontra-se disponível em: <https://doi.org/10.1039/d2ay01636b>).



### 4.3.1 Espectros de Fluorescência

Um exemplo observado da matriz excitação-emissão (EEM) de fluorescência para uma amostra de suco de uva integral é apresentado na Figura 35. No espectro, pode-se observar os espalhamentos Rayleigh e Raman e as intensidades de compostos químicos orgânicos que emitem radiação fluorescente. A escolha do valor de pH igual a 2 como condição do meio químico das amostras é baseada em trabalho anterior, em que os autores estabeleceram perfis de emissão para diferentes valores de pH investigados<sup>143</sup>. É possível visualizar mudanças na formação e intensidade dos espectros de fluorescência devido ao aumento do pH, no qual em pHs mais ácidos, há uma intensidade espectral relevante das espécies presentes, apresentando a melhor relação entre a intensidade do emitido sinal e o ruído instrumental.

**Figura 35.** Matriz excitação-emissão (EEM) de fluorescência para uma amostra de suco de uva integral.



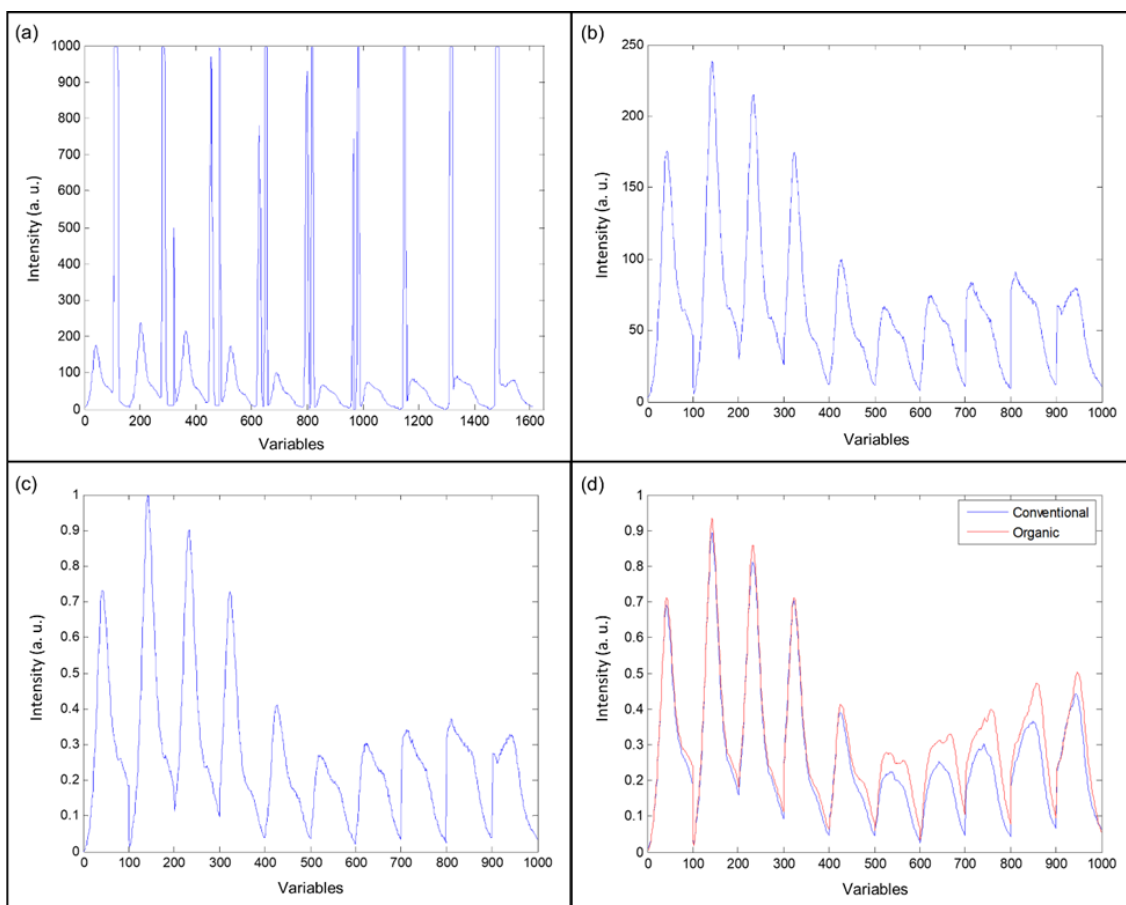
Fonte: elaborada pelo autor, 2022.

Os registros fluorescentes para cada amostra foram arranjados em uma matriz (excitação  $\times$  emissão para valor de pH = 2) de tamanho  $10 \times 161$ . Todas as matrizes foram desdobradas reduzindo a dimensionalidade dos dados de 2ª ordem para os dados de 1ª ordem. Este procedimento consiste em colocar cada espectro de emissão um ao lado do

outro para cada comprimento de onda de excitação, obtendo uma matriz de tamanho  $1 \times 1610$ . A importância de realizar o processo de desdobramento é facilitar a manipulação dos dados, reduzir informações irrelevantes ou redundantes, diminuir tempo computacional e espaço de armazenamento<sup>28</sup>.

A Figura 36 (a) mostra os perfis de excitação para os comprimentos de onda de emissão após o desdobramento. Para cada comprimento de excitação, os espectros foram cortados em 100 comprimentos de emissão, a fim de eliminar os sinais de 1ª e 2ª ordem dos espalhamentos Rayleigh e Raman, conforme mostrado na Figura 36 (b). A seguir, os dados foram normalizados como ilustrado na Figura 36 (c). Finalmente, a Figura 36 (d) apresenta os valores médios dos sinais dos dois conjuntos de amostras (convencional ou orgânica), em que é possível observar diferenças para as classes estudadas, o que foi um indício de que a construção de modelos de classificação supervisionada pode ser bem sucedida.

**Figura 36.** Perfis de emissão para os 10 comprimentos de onda de excitação de fluorescência após (a) desdobramento, (b) remoção dos espalhamentos Rayleigh e Raman, (c) normalização e (d) valores médios dos sinais por classe convencional ou orgânica.



Fonte: elaborada pelo autor, 2022.

### 4.3.2 Modelos LDA e PLS-DA

A construção dos modelos LDA e a seleção das variáveis mais relevantes foram realizadas utilizando os algoritmos ACO e GA. Os parâmetros utilizados para avaliar os diferentes modelos foram os seguintes: para o algoritmo ACO, o número de formigas foi 50, 100 e 200; número de colônias de 100, 200 e 500; taxa de formigas cegas de 0,25%, 0,35%, 0,50% e 0,70%; número máximo de variáveis selecionadas de 15 e 25; e taxa de evaporação de feromônio de 40% e 65%. Para o algoritmo GA, população de 50, 100 e 200; número de gerações de 100, 200 e 500; taxa de mutação de 5%, 10% e 15%; número máximo de variáveis selecionadas de 15, 25 e 40; e taxa de cruzamento de 40%, 60% e 80%. Com as combinações de parâmetros para cada modelo, elas foram realizadas em triplicata a fim de buscar a melhor solução e, de maneira geral, apresentaram alto índice de acerto na classificação das amostras, para ambas as classes estudadas.

O melhor modelo obtido para ACO-LDA teve as seguintes entradas: 100 para o número de formigas, 500 para o número de colônias, máximo de 15 variáveis selecionadas, taxa de formiga cega de 0,5% e taxa de feromônio evaporado de 65%. Para o GA-LDA, o modelo que apresentou o melhor resultado teve os seguintes critérios: 100 populações, 500 gerações, máximo de 15 variáveis, taxas de mutação e crossover de 10% e 40%, respectivamente. O modelo LDA empregando o algoritmo ACO apresentou desempenho ligeiramente superior ao GA-LDA com relação as figuras de mérito do conjunto de teste, porém ambos os modelos obtiveram alta acurácia para os conjuntos de treinamento e teste. O modelo PLS-DA foi aplicado a todo o conjunto de dados para comparar seu desempenho com os modelos LDA, e os resultados mostraram uma diminuição na precisão com perda considerável de seletividade para a classe de sucos de uva produzidos organicamente. As figuras de mérito, bem como o número de acertos e erros para cada um dos conjuntos de treinamento e teste, para cada algoritmo empregado são apresentados na Tabela 11.

**Tabela 11.** Figuras de mérito dos melhores modelos de classificação para dados de espectroscopia de fluorescência.

Classe Atribuída		Modelo ACO-LDA			
		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	11	1	6	0
	Orgânico	0	19	1	8

Sensibilidade (%)		91,7	100	100	88,9
Especificidade (%)		100	91,7	88,9	100
Acurácia (%)		95,8		94,4	
Modelo GA-LDA					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	12	0	6	0
	Orgânico	1	18	2	7
Sensibilidade (%)		100	94,7	100	77,8
Especificidade (%)		94,7	100	77,8	100
Acurácia (%)		97,4		88,9	
Modelo PLS-DA com 4 variáveis latentes					
Classe Atribuída		Treinamento		Teste	
		Convencional	Orgânico	Convencional	Orgânico
Matriz de Confusão	Convencional	10	2	5	1
	Orgânico	5	14	7	2
Sensibilidade (%)		83	74	83	78
Especificidade (%)		74	83	78	83
Acurácia (%)		77		80	

Fonte: elaborada pelo autor, 2022.

A Tabela 11 mostra que o melhor modelo ACO-LDA apresentou do conjunto de treinamento taxas de 91,7% de sensibilidade e 100% de seletividade para a classe convencional e taxas de 100% de sensibilidade e 91,7% de seletividade para a classe orgânica. O resultado obtido para a taxa de acurácia foi de 95,8% para as classes convencional e orgânica. Do conjunto de teste, obtiveram-se índices de 100% de sensibilidade e 88,9% de seletividade para a classe convencional e índices de 88,9% de sensibilidade e 100% de seletividade para a classe orgânica. Para ambas as classes, o resultado obtido para a taxa de acurácia foi de 94,4%.

O melhor modelo GA-LDA apresentou do conjunto de treinamento taxas de 100% de sensibilidade e 94,7% de seletividade para a classe convencional e taxas de 94,7% de sensibilidade e 100% de seletividade para a classe orgânica. Para as classes convencional e orgânica, o resultado obtido para a taxa de acurácia foi de 97,4%. Do conjunto de teste, obtiveram-se índices de 100% de sensibilidade e 77,8% de seletividade para a classe

convencional e índices de 77,8% de sensibilidade e 100% para a classe orgânica. O resultado obtido para a taxa de acurácia foi de 88,9% para as classes convencional e orgânica.

O modelo PLS-DA apresentou como resultado as taxas de 83% de sensibilidade e 74% de seletividade para a classe convencional e taxas de 74% de sensibilidade e 83% de seletividade para a classe orgânica do conjunto de treinamento. A taxa de acurácia para ambas as classes convencional e orgânica foi de 77%. Do conjunto de teste, obtiveram-se taxas de 83% de sensibilidade e 78% de seletividade para a classe convencional e taxas de 78% de sensibilidade e 83% de seletividade para a classe orgânica. A taxa de 80% de acurácia foi obtida para ambas as classes convencional e orgânica. Os resultados evidenciaram um desempenho inferior nas figuras de mérito do modelo PLS-DA em relação aos modelos LDA.

Gráficos de função discriminante para as duas classes, dos melhores modelos para ambos os algoritmos, são apresentados na Figura 37. As amostras de treinamento para os sucos produzidos no sistema convencional (círculos azuis) e teste (estrelas azuis) apresentaram valores negativos para a função discriminante e, por outro lado, as amostras de treinamento para os sucos produzidos no sistema orgânico (círculos vermelhos) e teste (estrelas vermelhas) apresentam função discriminante positiva.

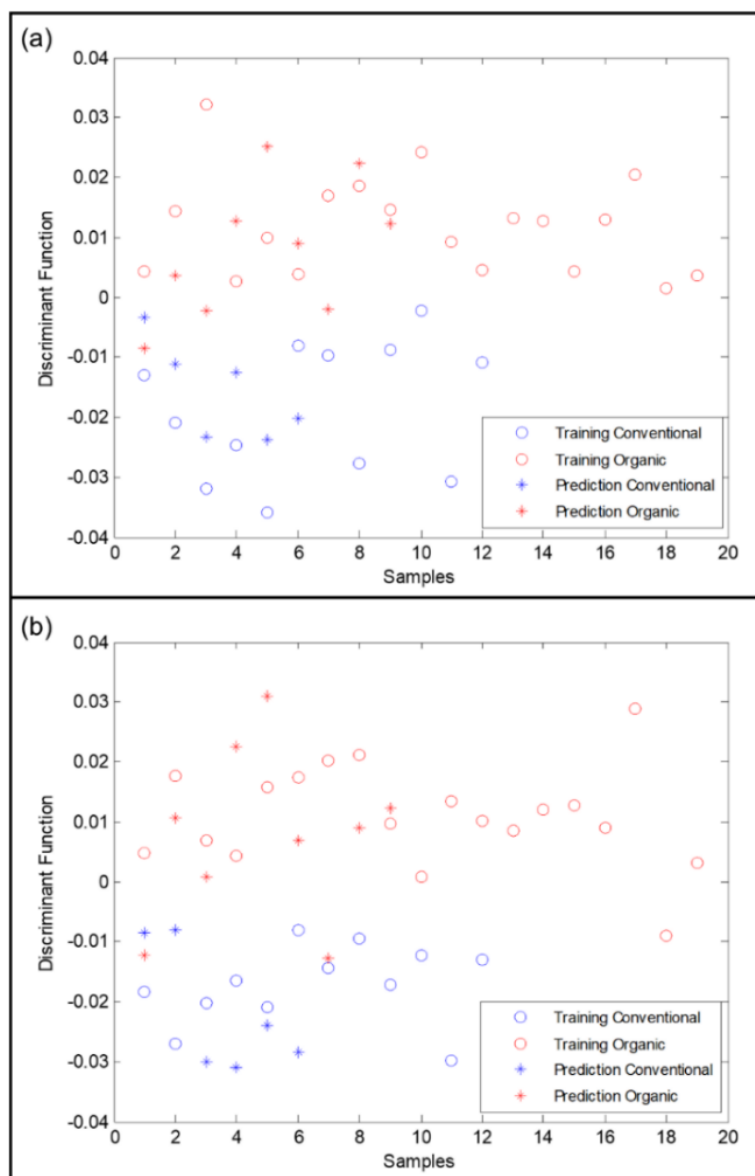
Os resultados mostraram que para o modelo ACO-LDA houve apenas uma amostra de treino da classe convencional classificada como orgânica e, também, uma amostra de teste da classe orgânica classificada como convencional. Para o modelo GA-LDA, uma amostra de treinamento da classe orgânica foi classificada como convencional e duas amostras de teste da classe orgânica foram classificadas como convencional.

Os sucos de uva apresentam um conjunto complexo e variado de substâncias químicas, possuindo baixo teor de proteínas e lipídios e sendo abundantes em outros metabólitos primários como carboidratos, sais minerais, ácidos orgânicos e vitaminas, e em metabólitos secundários, incluindo os compostos fenólicos, isto é, substâncias antioxidantes possuindo um anel aromático com um ou mais substitutos hidroxila, incluindo seus grupos funcionais. Estes compostos contribuem para as características sensoriais dos sucos, nomeadamente ao nível da cor, aroma, amargor, adstringência e estabilidade oxidativa<sup>182</sup>.

A composição química está ligada à matéria-prima utilizada, ao estado de maturação, ao modo de cultivo e aos tratamentos aplicados à uva<sup>19</sup>. Os sucos de uva apresentam sinais distintos de fluorescência com intensidade variável, de acordo com a

quantidade relativa de compostos químicos presentes, que também pode variar dependendo do sistema agrícola utilizado<sup>150</sup>.

**Figura 37.** Gráficos de função discriminante das amostras de treinamento e teste dos melhores modelos por (a) ACO-LDA e (b) GA-LDA.



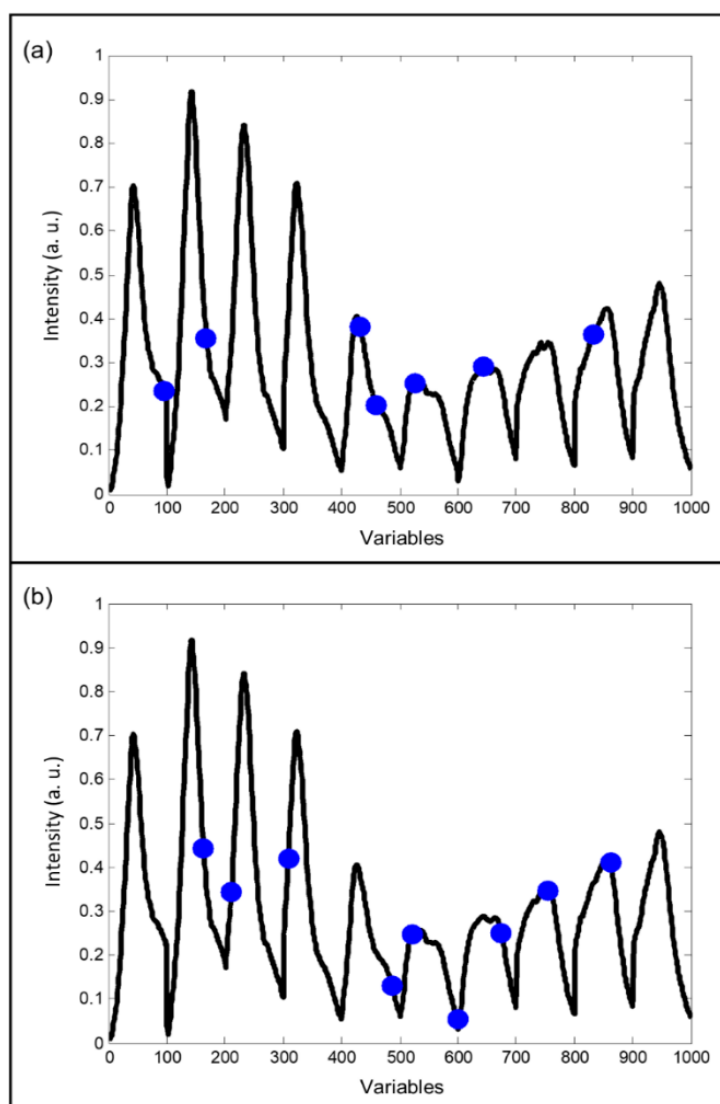
Fonte: elaborada pelo autor, 2022.

### 4.3.3 Seleção de Variáveis

A clara separação das duas classes de suco de uva integral pode ser explicada pelas variáveis mais relevantes selecionadas para os melhores modelos obtidos, identificadas como pontos azuis na Figura 37. Para o modelo ACO-LDA, as variáveis mais

representativas foram: 96, 168, 432, 460, 526, 644 e 832; e, para o modelo GA-LDA, as variáveis mais representativas foram: 163, 211, 309, 487, 522, 600, 675, 755 e 862. Essa seleção de variáveis se referem aos comprimentos de onda que representam, principalmente, os grupos de compostos fenólicos como flavonóides, antocianinas, catequinas monoméricas, taninos e ácidos fenólicos<sup>151</sup>.

**Figura 38.** Seleção das variáveis mais relevantes dos melhores modelos por (a) ACO-LDA e (b) GA-LDA.



Fonte: elaborada pelo autor, 2022.

A informação espectral está relacionada com compostos voláteis e polifenóis das uvas, que contêm sistemas  $\pi$ -conjugados dos grupos hidroxifenólicos<sup>183</sup>. Sinais característicos de pigmentos poliméricos e compostos de antocianinas também foram observados. O grupo das antocianinas inclui as antocianinas glucosidadas, aciladas, p-

cumaroiladas e os pigmentos poliméricos<sup>13,23</sup>. No entanto, a maioria das variáveis selecionadas para identificar as duas classes correspondem aos comprimentos de onda de excitação maiores, para os quais existem maiores intensidades médias para sucos de uva na classe orgânica, associadas a uma maior quantidade de compostos fenólicos conforme mostrado na Figura 35 (d).



## 5. CONCLUSÕES

No primeiro trabalho, uma nova metodologia para a diferenciação de sucos de uva integral produzidos no Brasil foi desenvolvida baseada na espectroscopia na região do visível e utilização da técnica multivariada supervisionada (GA-LDA). Os resultados obtidos denotaram a correta classificação das amostras de suco de uva de acordo com a técnica agrícola utilizada, convencional ou orgânica, com 100% de acerto na validação dos modelos e 92% na classificação das amostras externas. Destaca-se também o melhor desempenho do modelo GA-LDA em relação ao PLS-DA. O método proposto mostra-se relevante como opção no controle de qualidade, destacando-se por ser acessível e de fácil execução.

Na segunda pesquisa, a metodologia desenvolvida seguiu o conceito de Química Analítica Verde fundamentada nos princípios do desenvolvimento sustentável, apresentando como vantagens: o mínimo preparo de amostras, não utilizar solventes e reagentes químicos e, conseqüentemente, não gerar resíduos. Além disso, a análise espectroscópica no infravermelho médio, aliada à Quimiometria, apresenta-se como uma ferramenta atrativa devido às inúmeras vantagens da instrumentação como a rapidez nos resultados, simplicidade de operação, dados confiáveis e a precisão e confiabilidade dos algoritmos. O conjunto de dados FTIR se mostrou robusto devido a sua capacidade de ser aplicado em diferentes métodos de classificação supervisionados por análise discriminante, apresentando como resultados, modelos com excelentes índices de acurácia com 100% nos conjuntos de treinamento e de teste na predição das classes de suco de uva. Os principais sinais espectrais foram obtidos usando LDA em conjunto com os algoritmos GA, SW, ACO e SPA. Essa seleção permitiu distinguir regiões do espectro que funcionam como marcadores de grupos funcionais que permitem a identificação e categorização de amostras de suco de uva como convencional ou orgânica produzidos no Brasil. O método proposto é uma alternativa atrativa para o controle de qualidade a fim de avaliar fraudes em sucos de uva orgânicos, além das vantagens de ser ecologicamente correto e de simples aplicação.

No terceiro estudo, um novo método foi desenvolvido para classificar sucos de uva produzidos no Brasil com base em espectroscopia de fluorescência e Quimiometria. Os melhores resultados demonstraram eficiência na classificação de amostras de suco de uva de acordo com o processo de produção convencional ou orgânico, com até 97% para os modelos e 94% na predição dessas classes para amostras externas ao modelo. Os

melhores modelos ACO-LDA e GA-LDA apresentaram alta sensibilidade e especificidade, além de selecionar as variáveis mais significativas que explicam a separação das classes em questão. Além disso, a metodologia proposta revelou que os sucos de uva produzidos organicamente apresentam maiores teores de compostos fluorescentes nas condições estudadas. Desse modo, os dados de fluorescência podem ser aplicados como alternativa na identificação de sucos de uva integral, tornando este método viável para controle de qualidade em indústrias alimentícias a fim de auxiliar na descoberta de possíveis adulterações e/ou falsificações.

Portanto, ao se analisar os estudos desenvolvidos nessa tese, se conclui que os métodos analíticos desenvolvidos são eficientes e confiáveis e podem se tornar uma ferramenta valiosa na segurança alimentar para o controle de qualidade na autenticação de sucos de uva integral orgânicos. Dentre as soluções obtidas, a que se melhor destaca é a técnica da espectrometria na região do infravermelho médio associado a ferramenta quimiométrica GA-LDA, devido a maior facilidade de operação e rápido resultado do instrumento em conjunto com a simplicidade e eficiência dos algoritmos computacionais para a construção dos modelos e predição das classes das amostras. Assim, torna-se necessário esforços coletivos para o aprimoramento da legislação brasileira a fim de incluir esses novos métodos na avaliação da qualidade de sucos de uva.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS

<sup>1</sup> Pacheco, M.; Sgarbieri, V. Revisão: Alimentos funcionais fisiológicos. *Braz. J. Food Technol.*, 1999, 2, 7-19.

<sup>2</sup> Safraid, G.F., Portes, C.Z., Dantas, R.M., Batista, A.G. Profile of functional food consumer: identity and habits. *Braz. J. Food Technol.*, 2022, 25, e2021072.

<sup>3</sup> Shils, M.E.; Olson, J.A.; Shike, M. *Modern nutrition in health and disease*. 8ª ed., 2 v., Philadelphia, Lea and Febiger, 1994, 1646 p.

<sup>4</sup> Sang, J.; Sang, J.; Ma, Q.; Hou, X.; Li, C. Extraction optimization and identification of anthocyanins from *Nitraria tangutorun* Bobr. seed meal and establishment of a green analytical method of anthocyanins. *Food Chem.*, 2017, 218, 386-395.

<sup>5</sup> Baker, M.T.; Lu, P.; Parrella, J.A.; Leggette, H.R. Consumer acceptance toward functional foods: a scoping review. *Int. J. Environ. Res. Public Health*, 2022, 19, 1217.

<sup>6</sup> Granato, D.; Carrapeiro, M.M.; Fogliano, V.; Ruth, S.M. Effects of geographical origin, varietal and farming system on the chemical composition and functional properties of purple grape juices: a review. *Trends Food Sci. Technol.*, 2016, 52, 31-48.

<sup>7</sup> Toscano, L.T.; Silva, A.S.; Toscano, L.T.; Tavares, R.L.; Biasoto, A.C.T.; Camargo, A.C.; Silva, C.S.O.; Gonçalves, M.C.R.; Shahidi, F. Phenolics from purple grape juice increase serum antioxidant status and improve lipid profile and blood pressure in healthy adults under intense physical training. *J. Funct. Foods*, 2017, 33, 419-424.

<sup>8</sup> Zhou, K.; Raffoul, J.J. Potential anticancer properties of grape antioxidants. *J. Oncol.*, 2012, 2012, 1-8.

<sup>9</sup> Alaei, H.; Siahmard, Z.; Reisi, P. Chapter 86 - Red Grape Juice and Alzheimer's Disease. In: Martin, C.R.; Preedy, V.R. (Edts.). *Diet and Nutrition in Dementia and Cognitive Decline*. London, Academic Press, 2015, 935-942.

<sup>10</sup> Dani, C.; Dias, K.M.; Trevizol, L.; Bassôa, L.; Fraga, I.; Proença, I.C.T.; Pochmann, D.; Elsner, V.R. The impact of red grape juice (*Vitis labrusca*) consumption associated with physical training on oxidative stress, inflammatory and epigenetic modulation in healthy elderly women. *Physiol. Behav.*, 229, 2021, 113215.

<sup>11</sup> Lima, M.S.; Silani, I.S.V.; Toaldo, I.M.; Corrêa, L.C.; Biasoto, A.C.T.; Pereira, G.E.; Bordignon-Luiz, M.T.; Ninow, J.L. Phenolic compounds, organic acids and antioxidant activity of grape juices produced from new Brazilian varieties planted in the Northeast Region of Brazil. *Food Chem.*, 161, 2014, 94-103.

<sup>12</sup> Hosu, A.; Cristea, V.-M.; Cimpoiu, C. Analysis of total phenolic, flavonoids, anthocyanins and tannins content in Romanian red wines: prediction of antioxidant activities and classification of wines using artificial neural networks. *Food Chem.*, 150, 2014, 113-118.

<sup>13</sup> Padilha, C.V.S.; Miskinis, G.A.; Souza, M.E.A.O, Pereira, G.E.; Oliveira, D.; Bordignon-Luiz, M.T.; Lima, M.S. Rapid determination of flavonoids and phenolic acids in grape juices and wines by RP-HPLC/DAD: Method validation and characterization of commercial products of the new Brazilian varieties of grape. *Food Chem.*, 228, 2017, 106-115.

<sup>14</sup> Xu, W.; Liu, B.; Wang, C.; Kong, X. Organic cultivation of grape affects yeast succession and wine sensory quality during spontaneous fermentation. *LWT*, 120, 2020, 108894.

<sup>15</sup> Provost, C.; Pedneault, K. The organic vineyard as a balanced ecosystem: Improved organic grape management and impacts on wine quality. *Sci. Hortic.*, 208, 2016, 43-56.

<sup>16</sup> Chausali, N.; Saxena, J. Chapter 15 - Conventional versus organic farming: Nutrient status. In: Meena, V.S.; Meena, S.K.; Rakshit, A.; Stanley, J.; Srinivasarao, C. *Advances in Organic Farming: Agronomic Soil Management Practices*, Woodhead Publishing, 2021, 241-254.

<sup>17</sup> Filho, W.G.V. *Bebidas Não Alcoólicas: Ciência e Tecnologia*. 1<sup>a</sup> ed., 2 v., São Paulo, Blucher, 2010, 385 p.

- <sup>18</sup> Roldan, B.B. Influência do método de extração sobre a composição química de suco de uva bordô (*Vitis labrusca*). 2016. 57 f. Dissertação (Mestrado em Ciência e Tecnologia de Alimentos) - Instituto de Ciência e Tecnologia de Alimentos, Universidade Federal do Rio Grande do Sul, Porto Alegre.
- <sup>19</sup> Bresolin, B.; Gularte, M.A.; Manfroi, V. Exogenous water in grape juice obtained through the steam extraction method. *Rev. Bras. Tecnol. Agroind.*, 7, 2013, 922-933.
- <sup>20</sup> Rizzon, L.A.; Meneguzzo, J. Suco de uva - Coleção Agroindústria Familiar. Brasília, Embrapa Informação Tecnológica, 2007, 45 p.
- <sup>21</sup> Rizzon, L.A. Metodologia para análise de mosto e suco de uva. Brasília, Embrapa Informação Tecnológica, 2010, 78 p.
- <sup>22</sup> Sen, I.; Tokatli, F. Differentiation of wines with the use of combined data of UV-visible spectra and color characteristics. *J. Food Compos. Anal.*, 45, 2016, 101-107.
- <sup>23</sup> Aleixandre-Tudo, J.L.; Nieuwoudt, H.; Olivieri, A.; Aleixandre, J.L.; du Toit, W. Phenolic profiling of grapes, fermenting samples and wines using UV-Visible spectroscopy with chemometrics. *Food Control*, 85, 2018, 11-22.
- <sup>24</sup> Beltrame, K.K.; Gonçalves, T.R.; Gomes, S.T.M.; Matsushita, M.; Rutledge, D.N.; Março, P.H.; Valderrama, P. Digital images and independent components analysis in the determination of bioactive compounds from grape juice. *LWT*, 152, 2021, 112308.
- <sup>25</sup> Filho, P.A.C.; Chen, Y.; Cavin, C.; Galluzzo, R. Mid-infrared spectroscopy: Screening method for analysis of food adulterants in reconstituted skimmed milk powder. *Food Control*, 136, 2022, 108884.
- <sup>26</sup> Souza, R.R.; Fernandes, D.D.S.; Diniz, P.H.G.D. Honey authentication in terms of its adulteration with sugar syrups using UV-Vis spectroscopy and one-class classifiers. *Food Chem.*, 365, 2021, 130467.

- <sup>27</sup> He, M.; Chen, X.; Zhang, J.; Li, J.; Zhao, D.; Huang, Y.; Huo, D.; Luo, X.; Hou, C. Identification of liquors from the same brand based on ultraviolet, near-infrared and fluorescence spectroscopy combined with chemometrics. *Food Chem.*, 400, 2023, 134064.
- <sup>28</sup> Ferreira, M.M.C. *Quimiometria: conceitos, métodos e aplicações*. 1ª ed., Campinas, Unicamp, 2015, 493 p.
- <sup>29</sup> Aleixandre-Tudo, J.L.; Castello-Cogollos, L.; Aleixandre, J.L.; Aleixandre-Benavent, R. Chemometrics in food science and technology: A bibliometric study. *Chemometr. Intell. Lab. Syst.*, 222, 2022, 104514.
- <sup>30</sup> Moghaddam, H.N.; Tamiji, Z.; Lakeh, M.A.; Khoshayand, M.R.; Mahmoodi, M.H. Multivariate analysis of food fraud: a review of NIR based instruments in tandem with chemometrics. *J. Food Compos. Anal.*, 107, 2022, 104343.
- <sup>31</sup> Liu, Y.; Huang, J.; Li, M.; Chen, Y.; Cui, Q.; Lu, C.; Wang, Y.; Li, L.; Xu, Z.; Zhong, Y.; Ning, J. Rapid identification of the green tea geographical origin and processing month based on near-infrared hyperspectral imaging combined with chemometrics. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 267, 2022, 120537.
- <sup>32</sup> Fang, H.; Wu, H.-L.; Wang, T.; Long, W.-J.; Chen, A.-Q.; Ding, Y.-J.; Yu, R.-Q. Excitation-emission matrix fluorescence spectroscopy coupled with multi-way chemometric techniques for characterization and classification of Chinese lager beers. *Food Chem.*, 342, 2021, 128235.
- <sup>33</sup> Leão, P.C.S. Breve histórico da vitivinicultura e a sua evolução na região semiárida brasileira. *Anais da Academia Pernambucana de Ciência Agrônoma*, Recife, 7, 2010, 81-85.
- <sup>34</sup> Johnson, H. *A história do vinho*. São Paulo, Cia. das Letras, 1999, 546 p.
- <sup>35</sup> Fernandes, A. *Uso de Espectroscopia de Infravermelho com Transformada de Fourier (IVTF) e Quimiometria para Classificação de Vinho e Suco de Uva*. 2015. 39 f. Trabalho de Conclusão de Curso (Bacharelado em Química) - Universidade Tecnológica Federal do Paraná, Pato Branco, 2015.

- <sup>36</sup> Venquiaruto, L.D.; Dallago, R.M. Química das bebidas. Erechim, EdiFAPES, 2018.
- <sup>37</sup> De Rossi, G. Identificação e Análise do Sistema de Monitoramento de Fornecedores da Cadeia de Suprimentos do Setor Vitivinícola do Rio Grande do Sul: o Caso da Vinícola Miolo. 2009. 66 f. Trabalho de Conclusão de Curso (Especialização em Gestão de Operações Logísticas) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.
- <sup>38</sup> Bicas, T.C.; Fernandes, A.; Prasniewski, A.; Calegari, M.A.; Lima, V.A.; Oldoni, T.L.C. Uso da espectroscopia de infravermelho médio com transformada de Fourier (IV-TF) aliada à quimiometria para classificação de vinhos e suco de uva. *Braz. J. Food Res.*, 8, 2017, 89-97.
- <sup>39</sup> Food and Agriculture Organization (FAO). Production of grapes. 2020. Disponível em: [https://www.fao.org/faostat/en/#rankings/countries\\_by\\_commodity](https://www.fao.org/faostat/en/#rankings/countries_by_commodity). Acesso em: 06/10/2022.
- <sup>40</sup> International Organisation of Vine and Wine (OIV). Vine and Wine Outlook, Paris, France, 2012. Disponível em: <https://congress.oiv.int/public/medias/4524/oiv-vine-and-wine-outlook-2012-completfinal.pdf>. Acesso em: 06/10/2022.
- <sup>41</sup> Instituto Brasileiro de Geografia e Estatística (IBGE). Sidra. Banco de Tabelas Estatísticas. Disponível em: <https://sidra.ibge.gov.br/home/lspa/brasil>. Acesso em: 11/10/2022.
- <sup>42</sup> Mello, L.M.R.; Machado, C.A.E. Vitivinicultura brasileira: Panorama 2020. Comunicado Técnico 223. 1ª ed., Bento Gonçalves, Embrapa Uva e Vinho, 2021, 18 p.
- <sup>43</sup> Atlas Socioeconômico do RS. Uva e Maçã. Disponível em: <https://atlassocioeconomico.rs.gov.br/uva-e-maca>. Acesso em: 10/10/2022.
- <sup>44</sup> Spinelli, F.R. Validação de métodos para detecção da adição de suco maçã em sucos de uva. 2013. 81 f. Dissertação (Mestrado em Biotecnologia) - Universidade de Caxias do Sul, Caxias do Sul, 2013.

<sup>45</sup> Zanus, M.C. Embrapa Uva e Vinho. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/9103859/artigo-safra-da-uva-2016----o-que-esta-acontecendo>. Acesso em: 18/04/2023.

<sup>46</sup> RS. Secretaria da Agricultura, Pecuária e Desenvolvimento Rural. Disponível em: <https://www.agricultura.rs.gov.br/dados-uvvas-vinhos>. Acesso em: 11/10/2022.

<sup>47</sup> Mariani, C.M.; Henkes, J.A. Agricultura orgânica x agricultura convencional: soluções para minimizar o uso de insumos industrializados. *Rev. Gest. Sust. Ambient.*, Florianópolis, 3, 2014, 315-338.

<sup>48</sup> Vasconcellos, M.A. Agricultura Convencional. Disponível em: <https://www.infoescola.com/economia/agricultura-convencional/>. Acesso em: 08/10/2022.

<sup>49</sup> Le Champion, A.; Oury, F.-X.; Heumez, E.; Rolland, B. Conventional versus organic farming systems: dissecting comparisons to improve cereal organic breeding strategies. *Org. Agric.*, 10, 2020, 63-74.

<sup>50</sup> Sumberg, J.; Giller, K.E. What is ‘conventional’ agriculture? *Glob. Food Sec.*, 32, 2022, 100617.

<sup>51</sup> Brasil. Casa Civil. Lei nº 10.831 de 23 de dezembro de 2003. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2003/110.831.htm](http://www.planalto.gov.br/ccivil_03/leis/2003/110.831.htm). Acesso em: 10/10/2022.

<sup>52</sup> Damo, A. Algumas considerações sobre agroecologia e produção orgânica de alimentos. *Revista Ambiente & Educação*, 17, 2012, 139-154.

<sup>53</sup> Araguaia, M. Alimentos orgânicos. *Brasil Escola*. Disponível em: <https://brasilecola.uol.com.br/saude/alimentos-organicos.htm>. Acesso em: 07/10/2022.

<sup>54</sup> Reboita, P.S. Análise socioeconômica do perfil dos consumidores de produtos orgânicos no município de Rio Grande-RS. 2018. 43 f. Monografia de Conclusão de Curso (Bacharelado em Ciências Econômicas) - Universidade Federal do Rio Grande, Rio Grande, 2018.



<sup>55</sup> Melo, G.W.B.; Botton, M.; Garrido, L.R. Produção orgânica de uva para mesa. Documento 92. Bento Gonçalves, Embrapa Uva e Vinho, 2015, 35 p.

<sup>56</sup> Verdi, L.; Marta, A.D.; Falconi, F.; Orlandini, S.; Mancini, M. Comparison between organic and conventional farming systems using Life Cycle Assessment (LCA): A case study with an ancient wheat variety. *Eur. J. Agron.*, 141, 2022, 126638.

<sup>57</sup> Rizzon, L.A.; Manfroi, V.; Meneguzzo, J. Elaboração de suco de uva na propriedade vitícola. Bento Gonçalves, Embrapa Uva e Vinho, 1998, 24 p.

<sup>58</sup> Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (Sebrae). O cultivo e o mercado da uva. 2016. Disponível em: <https://www.sebrae.com.br/sites/PortalSebrae/artigos/o-cultivo-e-o-mercado-da-uva,ae8da5d3902e2410VgnVCM100000b272010aRCRD>. Acesso em 08/10/2022.

<sup>59</sup> Norberto, P.M. Sistemas de condução em videira: análise agrônômica e ecofisiológica. 2006. 118 f. Tese (Doutorado em Agronomia) - Universidade Federal de Lavras, Lavras, 2006.

<sup>60</sup> Empresa Brasileira de Pesquisa Agropecuária (Embrapa). Embrapa comercializa cultivares de porta enxerto de videira de qualidade superior. 2017. Disponível em: <https://www.embrapa.br/busca-de-noticias/-/noticia/24689196/embrapa-comercializa-cultivares-de-porta-enxerto-de-videira-de-qualidade-superior>. Acesso em 08/10/2022.

<sup>61</sup> Brasil. Ministério da Agricultura, Pecuária e Abastecimento. Decreto nº 6.871 de 4 de junho de 2009. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_Ato2007-2010/2009/Decreto/D6871.htm](https://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Decreto/D6871.htm). Acesso em: 07/10/2022.

<sup>62</sup> Brasil. Casa Civil. Lei nº 7.678 de 8 de novembro de 1988. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/Leis/1980-1988/L7678.htm](https://www.planalto.gov.br/ccivil_03/Leis/1980-1988/L7678.htm). Acesso em: 07/10/2022.

<sup>63</sup> Canossa, A.T.; Reinehr, J.; Bem, B.P.; Allenbandt, R.; Wurz, D.A.; Kretschmar, A.A. Composição química e análise sensorial do suco de uva elaborado com três variedades cultivadas em Lages - Santa Catarina. Revista da Jornada da Pós-Graduação e Pesquisa - Congrega, URCAMP, 2017.

<sup>64</sup> Brasil. Ministério da Agricultura, Pecuária e Abastecimento. Instrução Normativa nº 14, de 8 de fevereiro de 2018. Disponível em: <https://www.agricultura.rs.gov.br/upload/arquivos/202008/17121016-instrucao-normativa-14-de-2018-piqs-vinhos-e-derivados-da-uva-e-do-vinho.pdf>. Acesso em: 09/10/2022.

<sup>65</sup> Genova, G.; Tosetti, R.; Tonutti, P. Berry ripening, pre-processing and thermal treatments affect the phenolic composition and antioxidant capacity of grape (*Vitis vinifera* L.) juice. *J. Sci. Food Agric.*, 96, 2016, 664e671.

<sup>66</sup> Hopfer, H.; Nelson, J.; Collins, T.S.; Heymann, H.; Ebeler, S.E. The combined impact of vineyard origin and processing winery on the elemental profile of red wines. *Food Chem.*, 172, 2015, 486e496.

<sup>67</sup> Marzarotto, V., Capítulo 19 - Suco de Uva. In Filho, W.G.V. (Ed.). *Bebidas alcoólicas: ciência e tecnologia*. 1 v., São Paulo, Blucher, 2010, 359-384.

<sup>68</sup> Holler, F.J.; Skoog, D.A.; Crouch, S.R. *Princípios de Análise instrumental*. 6ª ed., Porto Alegre, Bookman, 2009, 1056 p.

<sup>69</sup> Harris, D.C. *Análise Química Quantitativa*. 6ª ed., Rio de Janeiro, LTC, 2005, 876 p.

<sup>70</sup> Instituto Adolfo Lutz. *Métodos químicos e físicos para análise de alimentos*. 3ª ed., 1 v., São Paulo, IMESP, 1985.

<sup>71</sup> Norma técnica NF EN 12630. Fruit and vegetables juices - Determination of glucose, fructose, sorbitol and sucrose contents - Method using high performance liquid chromatography. European Committee for Standardization, 1999, 15 p.

<sup>72</sup> Schieber, A.; Keller, P.; Carle, R. Determination of phenolic acids and flavonoids of apple and pear by high-performance liquid chromatography. *J. Chromatogr. A*, 910, 2001, 265-273.

<sup>73</sup> Riberéau-Gayon, P.; Gloria, Y.; Maujean, A.; Dubourdieu, D.; Miccio, M.T. *Tratado de Enología 2: Química del Vino Estabilización y Tratamientos*. Buenos Aires: Hemisferio Sur, 2008. 655 p.

<sup>74</sup> Pavia, D.L.; Lampman, G.M.; Kriz, G.S.; Vyvyan, J.R. *Introdução à Espectroscopia*. 1<sup>a</sup> ed., São Paulo, Cengage Learning, 2010, 716 p.

<sup>75</sup> Silverstein, R.M.; Webster, F.X.; Kiemle, D.J., *Identificação Espectrométrica de Compostos Orgânicos*. 7<sup>a</sup> ed., Rio de Janeiro, LTC, 2007, 490 p.

<sup>76</sup> Ronan, P. EM Spectrum, Wikimedia Commons, 2007. Disponível em: [https://commons.wikimedia.org/wiki/File:EM\\_spectrum.svg](https://commons.wikimedia.org/wiki/File:EM_spectrum.svg). Acesso em: 10/10/2022.

<sup>77</sup> Solomons, T.W.G.; Fryhle, C.B. *Química Orgânica*. 10<sup>a</sup> ed., 1 v., Rio de Janeiro, LTC, 2012, 644 p.

<sup>78</sup> Solomons, T.W.G.; Fryhle, C.B. *Química Orgânica*. 10<sup>a</sup> ed., 2 v., Rio de Janeiro, LTC, 2012, 613 p.

<sup>79</sup> Bruice, P.Y. *Química Orgânica*. 4<sup>a</sup> ed., 1 v., São Paulo, Pearson Prentice Hall, 2006, 547 p.

<sup>80</sup> Fernandes, D.D.S. *Espectroscopia UV-VIS para avaliação de biodiesel e misturas biodiesel/diesel*. 2013. 75 f. Dissertação (Mestrado em Ciências Agrárias) - Universidade Estadual da Paraíba, Campina Grande, 2013.

<sup>81</sup> Bonato, M., *Espectroscopia de infravermelho médio com transformada de Fourier aplicada ao controle de qualidade na produção de biodiesel*. 2016. 109 f. Dissertação (Mestrado em Engenharia Química) - Universidade Federal de Santa Catarina, Florianópolis, 2016.

- <sup>82</sup> Forato, L.A.; Filho, R.B.; Osiro, D.; Bicudo, T.C.; Colnago, L.A. A Espectroscopia na região do Infravermelho e algumas aplicações. Documento 51, São Carlos, Embrapa Instrumentação Agropecuária, 2010, 14 p.
- <sup>83</sup> Pasquini, C. Near Infrared Spectroscopy: fundamentals, practical aspects and analytical applications. *J. Braz. Chem. Soc.*, 14, 2003, 198-219.
- <sup>84</sup> Atkins, P.; De Paula, J. *Atkins' Physical Chemistry*. 8<sup>a</sup> ed., London, Oxford University Press, 2006, 1184 p.
- <sup>85</sup> Coelho, A.L. Desenvolvimento de procedimento de derivatização fotoquímica para dois glicocorticóides sintéticos (prednisolona e triancinolona acetono) visando à análise espectrofluorimétrica de formulações farmacêuticas. 2004. 119 f. Dissertação (Mestrado em Química) - Pontifícia Universidade Católica, Rio de Janeiro, 2004.
- <sup>86</sup> Universidade Federal de Juiz de Fora (UFJF). Espectrometria de Luminescência Molecular. Disponível em: <http://www.ufjf.br/nupis/files/2015/10/aula-9-fluoresc%C3%Aancia-parte-1.pdf>. Acesso em: 08/10/2022.
- <sup>87</sup> Cruz, J.F.M. Caracterização de gasolinas por espectroscopia FT-Raman. 2003. 215 f. Tese (Doutorado em Ciências-Química) - Pontifícia Universidade Católica, Rio de Janeiro, 2003.
- <sup>88</sup> Trevisan, M.G.; Poppi, R.J. Química Analítica de Processos. *Quím. Nova*, 29, 2006, 1065-1071.
- <sup>89</sup> Santana, F.B.; Souza, A.M.; Almeida, M.R.; Breikreitz, M.C.; Filgueiras, P.R.; Sena, M.M.; Poppi, R.J. Experimento didático de quimiometria para classificação de óleos vegetais comestíveis por espectroscopia no infravermelho médio combinado com análise discriminante por mínimos quadrados parciais: um tutorial, parte V. *Quím. Nova*, 43, 2020, 371-381.
- <sup>90</sup> Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.*, 7, 1936, 179-188.

- <sup>91</sup> Marini, F. Classification Methods in Chemometrics. *Curr. Anal. Chem.*, 6, 2009, 72-79.
- <sup>92</sup> Junior, E.R.S.; Ren, T.I.; Cavalcanti, G.D.C. Análise de Discriminante Linear. Disponível em: <https://www.cin.ufpe.br/~if699/NotasdeAula/14.LDA.pdf>. Acesso em: 12/10/2022.
- <sup>93</sup> Mohammadi, M.; Al-Azab, F.; Raahemi, B.; Richards, G.; Jaworska, N.; Smith, D.; Salle, S.; Blier, P.; Knott, V. Data mining EEG signals in depression for their diagnostic value, *BMC Medical Inform. Decis. Mak.*, 15, 2015, 108.
- <sup>94</sup> Yang, W.-H.; Daí, D.-Q. Two-Dimensional Maximum Margin Feature Extraction for Face Recognition. *IEEE Trans. Syst. Man Cybern. B Cybern.*, 39, 2009, 1002-1012.
- <sup>95</sup> Oliveri, P.; Simonetti, R. Chemometrics for Food Authenticity Applications. In Downey, G. (Ed.). *Advances in Food Authenticity Testing*. 1ª ed., Cambridge, Woodhead Publishing, 2016, 702-728.
- <sup>96</sup> Brereton, R.G.; Lloyd, G.R. Partial least squares discriminant analysis: taking the magic away. *J. Chemom.*, 48, 2014, 213-225.
- <sup>97</sup> Ali, A. Chapter 11 - Dimensionality Reduction. Wavy AI Research Foundation, 2019. Disponível em: <https://medium.com/machine-learning-researcher/dimensionality-reduction-pca-and-lda-6be91734f567>. Acesso em: 13/10/2022.
- <sup>98</sup> Moreira, E.D.T. Classificação de Cigarros Usando Espectrometria NIRR e Métodos Quimiométricos de Análise. 2007. 83 f. Dissertação (Mestrado em Química) – Universidade Federal da Paraíba, Campina Grande, 2007.
- <sup>99</sup> Silva, J.P.F.; Macedo, A.I. Análise Discriminante Linear, 2020. Disponível em: <https://lamfo-unb.github.io/2020/10/10/An%C3%A1lise-Discriminante-Linear/>. Acesso em: 12/10/2022.

- <sup>100</sup> Pontes, M.J.C.; Pereira, C.F.; Pimentel, M.F., Vasconcelos, F.V.C.; Silva, A.G.B. Screening analysis to detect adulteration in diesel/biodiesel blends using near infrared spectrometry and multivariate classification. *Talanta*, 85, 2011, 2159-2165.
- <sup>101</sup> Wang, S.; Liu, K.; Yu, X.; Wu, D.; He, Y. Application of hybrid image features for fast and non-invasive classification of raisin. *J. Food Eng.*, 109, 2012, 531-537.
- <sup>102</sup> Ballabio, D. & Consonni, V. Classification tools in chemistry. Part 1: linear models. PLS-DA. *Anal. Methods* 5, 3790 (2013).
- <sup>103</sup> Silva, V.A.G.; Talhavini, M.; Zacca, J.J.; Trindade, B.R.; Braga; J.W.B. Discrimination of Black Pen Inks on Writing Documents Using Visible Reflectance Spectroscopy and PLS-DA. *J. Braz. Chem. Soc.*, 25, 2014, 1552-1564.
- <sup>104</sup> Gorski, Ł.; Sordón, W.; Ciepiela, F.; Kubiak, W.W.; Jakubowska, M. Voltammetric classification of ciders with PLS-DA. *Talanta*, 146, 2016, 231-235.
- <sup>105</sup> Medina, S.; Pereira, J.A.; Silva, P.; Perestrelo, R.; Câmara, J.S. Food fingerprints – A valuable tool to monitor food authenticity and safety. *Food Chem.*, 278, 2019, 144-162.
- <sup>106</sup> Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E.; Tatham, R.L. *Análise multivariada de dados*. 6<sup>a</sup> ed., Porto Alegre, Bookman, 2009, 682 p.
- <sup>107</sup> Pontes, M.J.C. Algoritmo das projeções sucessivas para a seleção de variáveis espectrais em problemas de classificação. 2009. 123 f. Tese (Doutorado em Química) - Universidade Federal da Paraíba, João Pessoa, 2009.
- <sup>108</sup> Soares, S.F.C. Um novo critério para seleção de variáveis usando o Algoritmo das Projeções Sucessivas. 2010. 107 f. Dissertação (Mestrado em Química) - Universidade Federal da Paraíba, João Pessoa, 2010.
- <sup>109</sup> Silva, W.R.; Paula, L.C.M.; Soares, A.S.; Coelho, C.J. Algoritmo do morcego para seleção de variáveis em calibração multivariada. *Rev. Arithmós*, 1, 2019, 13-17.

- <sup>110</sup> Whelan, G.; Sarmiento, R.; Sprenger, J. Universal-deterministic and probabilistic hypotheses in operations management research: a discussion paper. *Prod. Plan. Control*, 29, 2018, 1306-1320.
- <sup>111</sup> Lessio, F.; Alma, A. Models applied to grapevine pests: a review. *Insects*, 12, 2021, 169.
- <sup>112</sup> Holland, J.H. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MIT Press, 1992, 232 p.
- <sup>113</sup> Lacerda, E.G.M.; Carvalho, A.C.P.L.F. *Introdução aos algoritmos genéticos. Sistemas inteligentes: aplicações a recursos hídricos e sistemas ambientais*. Porto Alegre, Ed. Universidade/UFRG/ABRH, 1999.
- <sup>114</sup> Miaw, C.S.W; Sena, M.M.; Souza, S.V.C.; Ruisanchez, I.; Callao, M.P. Variable selection for multivariate classification aiming to detect individual adulterants and their blends in grape nectars. *Talanta*, 190, 2018 55-61.
- <sup>115</sup> Obitko, M. *Introdução aos Algoritmos Genéticos*. Disponível em: <https://www.obitko.com/tutorials/genetic-algorithms/portuguese/about.php>. Acesso em: 12/10/2022.
- <sup>116</sup> Carvalho, A.P.L.F. *Algoritmos Genéticos*. Disponível em: <https://sites.icmc.usp.br/andre/research/genetic/>. Acesso em: 13/10/2022.
- <sup>117</sup> Leardi, R. Genetic algorithms in chemometrics and chemistry: a review. *J. Chemom.*, 15, 2001, 559-569.
- <sup>118</sup> Filho, P.A.C.; Poppi, R.J. Algoritmo genético em química. *Quím. Nova*, 23, 1999, 405-411.
- <sup>119</sup> Kim, D.H.; Abraham, A.; Cho, J.H. A hybrid genetic algorithm and bacterial foraging approach for global optimization. *Inf. Sci.*, 177, 2007, 3918-3937.

- <sup>120</sup> Izquierdo, R.C. Projeto de formação de células de manufatura através da utilização de algoritmos genéticos. 2013. 93 f. Dissertação (Mestrado em Engenharia Mecânica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013.
- <sup>121</sup> Dorigo, M.; Gambardella, L.M. Ant colonies for the travelling salesman problem. *Biosystems*, 43, 1997, 73-81.
- <sup>122</sup> Allegrini, F.; Olivieri, A.C. A new and efficient variable selection algorithm based on ant colony optimization. Applications to near infrared spectroscopy/partial least-squares analysis. *Anal. Chim. Acta*, 699, 2011, 18-25.
- <sup>123</sup> Dorigo, M.; Stützle, T. *Ant Colony Optimization*. 1<sup>a</sup> ed., Cambridge, Bradford Book, 2004, 319 p.
- <sup>124</sup> Santos, R.L.; Leal, J.E. Solução de um problema de roteirização com janelas de tempo através de um algoritmo de múltiplas colônias de formigas. *Transportes*, 15, 2007, 5-16.
- <sup>125</sup> Ferreira, R.P.; Sassi, R.J.; Costa, F.M.; Ferreira, A.; Souza, A.S. Aplicando o algoritmo de otimização por colônia de formigas e os mapas auto-organizáveis de Kohonen na roteirização e programação de veículos. Disponível em: <https://www.marinha.mil.br/spolm/sites/www.marinha.mil.br/spolm/files/102807.pdf>. Acesso em: 14/10/2022.
- <sup>126</sup> Dorigo, M.; Gambardella, L.M. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Trans. Evol. Comput.*, 1, 1997, 53-66.
- <sup>127</sup> Zareef, M.; Chen, Q.; Ouyang, Q.; Arslan, M.; Hassan, M.M.; Ahmad, W.; Viswadevarayalu, A.; Wang, P.; Ancheng, W. Rapid screening of phenolic compounds in congou black tea (*Camellia sinensis*) during in vitro fermentation process using portable spectral analytical system coupled chemometrics. *J. Food Process. Preserv.*, 43, 2019, e13996.
- <sup>128</sup> Hu, L.; Yin, C.; Ma, S.; Liu, Z. Rapid detection of three quality parameters and classification of wine based on Vis-NIR spectroscopy with wavelength selection by ACO and CARS algorithms. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 205, 2018, 574-581.



- <sup>129</sup> Araújo, M.C.U.; Saldanha, T.C.B.; Galvão, R.K.H.; Yoneyama, T.; Chame, H.C.; Visani, V. The successive projections algorithm for variable selection in spectroscopic multicomponent analysis. *Chemometr. Intell. Lab. Syst.*, 57, 2001, 65-73.
- <sup>130</sup> Galvão, R.K.H.; Araújo, M.C.U. Linear regression modelling: variable selection. In: Tauler, R.; Walczak, B.; Brown, S. (Eds.). *Comprehensive chemometrics: chemical and biochemical data analysis*, 1<sup>a</sup> ed., 3 v., London, Elsevier, 2009, 233-283.
- <sup>131</sup> Silva, G.W.B.; Gomes, A.A.; Silva, P.; Costa, G.B.; Fernandes, D.D.S.; Fontes, M.M.; Veras, G. Biodiesel/diesel blends classification with respect to base oil using NIR spectrometry and chemometrics tools. *J. Am. Oil Chem. Soc.*, 89, 2012, 1165-1171.
- <sup>132</sup> Oliveira, V.; Knapic, S.; Pereira, H. Classification modeling based on surface porosity for the grading of natural cork stoppers for quality wines. *Food Bioprod. Process.*, 93, 2015, 69-76.
- <sup>133</sup> Alves, M.F.; Lotufo, A.D.P.; Lopes, M.L.M. Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas. *Proc. Ser. Braz. Soc. Appl. Comput. Math*, 1, 2013, 010144.
- <sup>134</sup> Draper, N.R.; Smith, H. *Applied Regression Analysis*. 3<sup>a</sup> ed., New York, John Wiley & Sons, 1998, 736 p.
- <sup>135</sup> Moreira, E.D.T.; Pontes, M.J.C; Galvão, R.K.H.; Araújo, M.C.U. Near infrared reflectance spectrometry classification of cigarettes using the successive projections algorithm for variable selection. *Talanta*, 79, 2009, 1260-1264.
- <sup>136</sup> Soares, S.F.C.; Gomes, A.A.; Filho, A.R.G.; Araújo, M.C.U.; Galvão, R.K.H. The successive projections algorithm. *Trends Anal. Chem.*, 42, 2013, 84-96.
- <sup>137</sup> Ghasemi-Varnamkhasti, M.; Mohtasebi, S.S.; Rodriguez-Mendez, M.L.; Gomes, A.A.; Araújo, M.C.U.; Galvão, R.K.H. Screening analysis of beer ageing using near infrared spectroscopy and the successive projections algorithm for variable selection. *Talanta*, 89, 2012, 286-291.

<sup>138</sup> Xu, M.; Sun, J.; Zhou, X.; Tang, N.; Shen, J.; Wu, X. Research on nondestructive identification of grape varieties based on EEMD-DWT and hyperspectral image. *J. Food Sci.*, 86, 2021, 2011-2023.

<sup>139</sup> Pontes, M.J.C.; Galvão, R.K.H.; Araújo, M.C.U.; Moreira, P.N.T.; Neto, O.D.P.; José, G.E.; Saldanha, T.C.B. The successive projections algorithm for spectral variable selection in classification problems. *Chemometr. Intell. Lab. Syst.*, 78, 2005, 11-18.

<sup>140</sup> Rech, A.M. Avaliação da espectroscopia no infravermelho combinada a análise multivariada para caracterização e controle de qualidade de bebidas saborizadas a base de soja. 2014. 47 f. Trabalho de Conclusão de Curso (Química Industrial) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014.

<sup>141</sup> Pasquini, C. Espectroscopia no Infravermelho Próximo (NIR): Módulo II - Métodos de Análise e Regressão Multivariada: Teoria e Prática. 2011. Disponível em: [http://www.gia.iqm.unicamp.br/artigospdfematerialcursos/Material%20didatico/Curso\\_QP812\\_2011\\_Parte%20II.pdf](http://www.gia.iqm.unicamp.br/artigospdfematerialcursos/Material%20didatico/Curso_QP812_2011_Parte%20II.pdf). Acesso em: 14/10/2022.

<sup>142</sup> Guerra, C.C.; Bitarelo, H.; Ben, R.L.; Marin, A. Sistema para elaboração de suco de uva integral em pequenos volumes: Suquificador integral. Documento 96. Bento Gonçalves, Embrapa Uva e Vinho, 2016, 32 p.

<sup>143</sup> Lenardon, L.V.; Böck, F.C.; Schneider, M.P.; Reis, N.A.P.C.; Silva, L.F.; de Souza, K.C.M.; Guerra, C.C.; Gomes, A.A.; Bergold, A.M.; Ferrão, M.F. Geographical origin authentication of southern Brazilian red wines by means of EEM-pH four-way data modelling coupled with one class classification approach. *Food Chem.*, 362, 2021, 130087.

<sup>144</sup> Grasel, F.S; Ferrão, M.F. A rapid and non-invasive method for the classification of natural tannin extracts by near-infrared spectroscopy and PLS-DA. *Anal. Methods.*, 8, 2016, 644-649.

- <sup>145</sup> Chen, W.; Chen, H.; Feng, Q.; Mo, L.; Hong, S. A hybrid optimization method for sample partitioning in near-infrared analysis. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, 248, 2021, 119182.
- <sup>146</sup> Li, X.; Fan, P.; Li, Z.; Chen, G.; Qiu, H.; Hou, G. Soil Classification Based on Deep Learning Algorithm and Visible Near-Infrared Spectroscopy. *J. Spectrosc.*, 2021, 2021, 1508267.
- <sup>147</sup> Gomes, A.A.; Azcarate, S.M.; Diniz, P.H.G.D.; Fernandes, D.D.S.; Veras, G. Variable selection in the chemometric treatment of food data: a tutorial review. *Food Chem.*, 370, 2022, 131072.
- <sup>148</sup> Oliveri, P.; Downey, G. Multivariate class modeling for the verification of food-authenticity claims. *Trends Anal. Chem.*, 35, 2012, 74-86.
- <sup>149</sup> Liu, Y.; Dang, B.; Li, Y.; Lin, H.; Ma, H. Applications of Savitzky-Golay filter of seismic random noise reduction. *Acta Geophys.*, 64, 2016, 101-124.
- <sup>150</sup> Grønder-Pedersen, L.; Rasmussen, S.E.; Bügel, S.; Jørgensen, L.V.; Dragsted, L.O.; Gundersen, V.; Sandström, B. Effect of diets based on foods from conventional versus organic production on intake and excretion of flavonoids and markers of antioxidative defense in humans. *J. Agric. Food Chem.*, 51, 2003, 5671-5676.
- <sup>151</sup> Sádecká, J.; Jakubíková, M. Varietal classification of white wines by fluorescence spectroscopy. *J. Food Sci. Technol.*, 57, 2020, 2545-2553.
- <sup>152</sup> Cozzolino, D.; Kwiatkowski, M.J.; Parker, M.; Cynkar, W.U.; Damberg, R.G.; Gishen, M.; Herderich, M.J. Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Anal. Chim. Acta*, 513, 2004, 73-80.
- <sup>153</sup> Levi, M.A.B.; Scarminio, I.S.; Poppi, R.J.; Trevisan, M.G. Three-way chemometric method study and UV-vis absorbance for the study of simultaneous degradation of anthocyanins in flowers of the *Hibiscus rosa-sinensis* species. *Talanta*, 62, 2004, 299-305.

- <sup>154</sup> Garcia, A.J.C.; Reis, B.F. Instrumentation and automated photometric titration: Procedure for total acidity determination in red wine employing a multicommutated flow system. *J. Autom. Methods Manag. Chem.*, 83247, 2006, 1-8.
- <sup>155</sup> Korzuchowska, A.; Wilska-Jeszka, J. Anthocyanins and chlorogenic acid copigmentation: influence on the colour of strawberry and chokeberry juices. *Z. Lebensm. Unter. Forsch.*, 203, 1996, 38-42.
- <sup>156</sup> Fernandes, A.; Oliveira, J.; Teixeira, N.; Mateus, N.; Freitas, V. A review of the current knowledge of red wine colour. *OENO One*, 51, 2017, 1-21.
- <sup>157</sup> Ferreiro-González, M.; Ruiz-Rodríguez, A.; Barbero, G.F.; Ayuso, J.; Alvarez, J.A.; Palma, M.; Barroso, C.G. FT-IR, Vis spectroscopy, color and multivariate analysis for the control of ageing processes in distinctive Spanish wines. *Food Chem.*, 277, 2019, 6-11.
- <sup>158</sup> Figueiredo-Gonzalez, M.; Cancho-Grande, B.; Simal-Gandara, J. Effects on colour and phenolic composition of sugar concentration processes in dried-on- or dried-off-vine grapes and their aged or not natural sweet wines. *Trends Food Sci. Technol.*, 31, 2013, 36-54.
- <sup>159</sup> Sádecká, J.; Jakubíková, M. Classification of Tokaj wines by ultraviolet-visible spectroscopy. *Food Anal. Methods*, 15, 2021, 56-66.
- <sup>160</sup> Marçõ, P.H.; Poppi, R.J.; Scarminio, I.S. Procedimentos analíticos para identificação de antocianinas presentes em extratos naturais. *Química Nova*, 31, 2008, 1218-1223.
- <sup>161</sup> Lee, J.; Durst, R.W.; Wrolstad, R.E. Determination of total monomeric anthocyanin pigment content of fruit juices, beverages, natural colorants, and wines by the pH differential method: Collaborative study. *J. AOAC Int.*, 88, 2005, 1269-1278.
- <sup>162</sup> Alcalde-Eon, C.; Escribano-Bailón, M.T.; Santos-Buelga, C.; Rivas-Gonzalo, J.C. Changes in the detailed pigment composition of red wine during maturity and ageing: A comprehensive study. *Anal. Chim. Acta*, 563, 2006, 238-254.

- <sup>163</sup> He, F.; Liang, N.-N.; Mu, L.; Pan, Q.-H.; Wang, J.; Reeves, M. J.; Duan, C.-Q. Anthocyanins and their variation in red wines II. Anthocyanin derived pigments and their color evolution. *Molecules*, 17, 2012, 1483-1519.
- <sup>164</sup> Cabrita, M.J.; Ricardo-da-Silva, J.; Laureano, O. Os compostos polifenólicos das uvas e dos vinhos. I Seminário Internacional de Vitivinicultura. Universidad Técnica de Lisboa, 2003, 61-102.
- <sup>165</sup> Junges, C.H.; Guerra, C.C.; Reis, N.A.P.C.; Gomes, A.A.; Diogo, F.S.; Ferrão, M.F. Grape juice classification with respect agricultural production system by means of visible spectroscopy chemometric assisted. *J. Food Compos. Anal.*, 114, 2022, 104793.
- <sup>166</sup> Kościelniak, P. White calibration methods based on gradient techniques as a modern approach to quantitative flow analysis. *Trends Anal. Chem.*, 157, 2022, 116758.
- <sup>167</sup> Borges, E.M.; Volmer, D.A.; Brandelero, E.; Gelinski, J.M.L.N.; Gallimberti, M.; Barbosa, F. Monitoring the authenticity of organic grape juice via chemometric analysis of elemental data. *Food Anal. Methods*, 9, 2016, 362-369.
- <sup>168</sup> Margraf, T.; Santos, T.N.T.; de Andrade, E.F.; van Ruth, S.M.; Granato, D. Effects of geographical origin, variety and farming system on the chemical markers and in vitro antioxidant capacity of Brazilian purple grape juices. *Food Res. Int.*, 82, 2016, 145-155.
- <sup>169</sup> Granato, D.; Koot, A.; van Ruth, S.M. Geographical provenancing of purple grape juices from different farming systems by proton transfer reaction mass spectrometry using supervised statistical techniques. *J. Sci. Food Agric.*, 95, 2015, 2668-2677.
- <sup>170</sup> Wongravee, K.; Ishigaki, M.; Ozaki, Y. Chapter 10 - Chemometrics as a Green Analytical Tool. In: Garrigues, S.; de la Guardia, M. (Edts.). *Challenges in Green Analytical Chemistry*. 2<sup>a</sup> ed., London, RSC, 2020.
- <sup>171</sup> Kościelniak, P. White calibration methods based on gradient techniques as a modern approach to quantitative flow analysis. *Trends Anal. Chem.*, 157, 116758, 2022.

<sup>172</sup> Anastas, P.T.; Warner, J.C. *Green Chemistry: theory and practice*. 1<sup>a</sup> ed., New York, Oxford University Press, 1998.

<sup>173</sup> Colthup, N.B.; Daly, L.H.; Wiberley, S.E. *Introduction to infrared and Raman spectroscopy*. 3<sup>a</sup> ed., California, Academic Press, 2012.

<sup>174</sup> Cozzolino, D. The role of visible and infrared spectroscopy combined with chemometrics to measure phenolic compounds in grape and wine samples. *Molecules*, 20, 2015, 726-737.

<sup>175</sup> Miaw, C.S.W.; Sena, M.M.; de Souza, S.V.C.; Callao, M.P.; Ruisachez, I. Detection of adulterants in grape nectars by attenuated total reflectance Fourier-transform mid-infrared spectroscopy and multivariate classification strategies. *Food Chem.*, 266, 2018, 254-261.

<sup>176</sup> Shah, N.; Cynkar, W.; Smith, P.; Cozzolino, D. Use of attenuated total reflectance midinfrared for rapid and real-time analysis of compositional parameters in commercial white grape juice. *J. Agric. Food Chem.*, 58, 2010, 3279-3283.

<sup>177</sup> Parpinello, G.P.; Ricci, A.; Arapitsas, P.; Curioni, A.; Moio, L.; Segade, S.R.; Ugliano, M.; Versari, A. Multivariate characterisation of Italian monovarietal red wines using MIR spectroscopy. *Oeno One*, 53, 2019, 741-751.

<sup>178</sup> Shen, F.; Wu, Q.; Su, A.; Tang, P.; Shao, X.; Liu, B. Detection of adulteration in freshly squeezed orange juice by electronic nose and infrared spectroscopy. *Czech J. Food Sci.*, 34, 2016, 224-232.

<sup>179</sup> Ferrer-Gallego, R.; Rodríguez-Pulido, F.J.; Toci, A.T.; García-Estevez, I. Phenolic composition, quality and authenticity of grapes and wines by vibrational spectroscopy. *Food Rev. Int.*, 38, 2022, 884-912.

<sup>180</sup> Mditshwa, A.; Magwaza, L.S.; Tesfay, S.Z.; Mbili, N. Postharvest quality and composition of organically and conventionally produced fruits: a review. *Sci. Hortic.*, 216, 2017, 148-159.

<sup>181</sup> Rahman, S.M.E.; Mele, M.A.; Lee, Y.-T.; Islam, M.Z. Consumer preference, quality, and safety of organic and conventional fresh fruits, vegetables, and cereals. *Foods*, 10, 2021, 105.

<sup>182</sup> Cosme, F.; Gonçalves, B.; Ines, A.; Jordão, A.M.; Vilela, A. Chapter 9 - Grape and wine metabolites: biotechnological approaches to improve wine quality. In: Morata, A.; Loira, I. (Edts.). *Grape and Wine Biotechnology*. London, IntechOpen, 2016, 470.

<sup>183</sup> Geaña, E.I.; Ciucure, C.T.; Apetrei, C.; Artem, V. Application of spectroscopic UV-Vis and FT-IR screening techniques coupled with multivariate statistical analysis for red wine authentication: varietal and vintage year discrimination. *Molecules*, 24, 2019, 4166.