

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**  
**FACULDADE DE FARMÁCIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS**  
**FARMACÊUTICAS**

Novas Metodologias para Determinação de Origem de Vinhos Gaúchos  
Empregando Espectrometria por Fluorescência Associada a Quimiometria

LAYANE LENARDON VINCIGUERRA

Porto Alegre - RS

2021



**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL**  
**FACULDADE DE FARMÁCIA**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS**  
**FARMACÊUTICAS**

Novas Metodologias para Determinação de Origem de Vinhos Gaúchos  
Empregando Espectrometria por Fluorescência Associada a Quimiometria

Tese apresentada por **Layane Lenardon**  
**Vinciguerra** para a obtenção do grau de Doutora  
em Ciências Farmacêuticas

**Orientadora:** Prof<sup>a</sup>. Dra. Ana Maria Bergold

**Coorientador:** Prof. Dr. Marco Flôres Ferrão

Porto Alegre -RS

2021

Tese apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas, em nível de Doutorado Acadêmico da Faculdade de Farmácia da Universidade Federal do Rio Grande do Sul e aprovada em 30 de março de 2021, pela Banca Examinadora constituída por:

Prof. Dr. Adilson Ben Da Costa  
Universidade de Santa Cruz do Sul

Prof<sup>a</sup>. Dra. Aline Rigon Zimmer  
Universidade Federal do Rio Grande do Sul

Dr. Marcelo Caetano Alexandre Marcelo  
Souza Cruz (BAT)

Prof. Dr. Pedro Eduardo Froehlich  
Universidade Federal do Rio Grande do Sul

#### CIP - Catalogação na Publicação

Vinciguerra, Layane Lenardon  
Novas Metodologias para Determinação de Origem de Vinhos Gaúchos Empregando Espectrometria por Fluorescência Associada a Quimiometria / Layane Lenardon Vinciguerra. -- 2021.  
156 f.  
Orientadora: Ana Maria Bergold.  
  
Coorientador: Marco Flôres Ferrão.  
  
Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Faculdade de Farmácia, Programa de Pós-Graduação em Ciências Farmacêuticas, Porto Alegre, BR-RS, 2021.  
  
1. Espectrometria de Fluorescência. 2. Vinho tinto. 3. Origem Geográfica. 4. Quimiometria. I. Bergold, Ana Maria, orient. II. Ferrão, Marco Flôres, coorient. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

## **AGRADECIMENTOS**

A Deus por tudo.

Ao Programa de Pós-graduação em Ciências Farmacêuticas (PPGCF) da Universidade Federal do Rio Grande do Sul (UFRGS) pela qualidade e oportunidade do mestrado e doutorado.

Ao meu marido, Robson Willians Vinciguerra, pelo amor, carinho, companheirismo, motivação, ensinamentos, paciência, por sempre acreditar no meu potencial e buscar o melhor para mim durante todos esses anos juntos.

A Prof<sup>ª</sup>. Dra. Ana Maria Bergold, pela orientação e oportunidade concedida, pelo exemplo de profissionalismo, pelas oportunidades de crescimento profissional e pela paciência e amizade.

Ao Prof. Dr. Marco Flôres Ferrão, pela coorientação deste trabalho, pela paciência e disponibilização do laboratório para o desenvolvimento do mesmo, pelos ensinamentos e dedicação, pela amizade e convívio durante todos esses anos de trabalho junto.

Ao Prof. Dr. Adriano Gomes de Araújo, disponibilidade de coorientação prestada durante esse trabalho, pelas ideias, pela paciência, amizade e ensinamentos.

A Dr. Celito Crivellaro Guerra, pela doação dos vinhos, disponibilidade, pelas visitas e toda atenção prestada para o melhoramento e execução desse trabalho.

Aos colegas do laboratório LAQIA, em especial, a Fernanda Carla Böck, pela amizade, paciência, dedicação, pelas ideias, pela parceria durante todo o doutorado e agradeço aos demais colegas que contribuíram com seus ensinamentos.

A toda equipe da EMBRAPA Uva e Vinhos de Bento Gonçalves pela doação de materiais, pelos equipamentos e todo suporte técnico prestado.

Ao Instituto de Química da Universidade Federal do Rio Grande do Sul pela disponibilidade dos equipamentos e laboratórios.

Enfim, agradeço a todas as pessoas que contribuíram de alguma forma para a realização deste trabalho.



## RESUMO

Novas metodologias por espectrofluorimetria acoplada às ferramentas quimiométricas foram desenvolvidas, visando a classificação de amostras de vinho tinto varietais produzidas em duas regiões do Rio Grande do Sul, considerando a sua origem geográfica. Como as propriedades fluorescentes de um composto químico podem variar em função do pH, e as mudanças estruturais resultantes do pH induzem alterações significativas nos seus espectros de fluorescência, foi possível separar e identificar também diferentes compostos químicos. Neste trabalho, foram analisadas 53 amostras de vinho tinto da região da Serra Gaúcha e 20 da região da Campanha, contemplando 10 variedades de uvas. O sinal de fluorescência registrado corresponde a nove matrizes de emissão (51 variáveis) de excitação (12 variáveis) (EEM) registradas em diferentes pH (3 até 11) gerando assim a matriz de dados representando os dados de ordem superior. Estes foram tratados pela Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes (MCR-ALS) e one class method Data Driven Soft Independent Modelling of Class Analogy (DD-SIMCA) para construção dos modelos de classificação. Ainda foram selecionados dois pHs (3 e 7) bem como a fusão dos dados destes, para representar dados de 1ª ordem, que foram explorados por meio da aplicação dos algoritmos (ACO, GA e SW) para seleção de variáveis, para auxiliar no reconhecimento dos vinhos empregando análise discriminante linear. Como resultado, observou-se que a predição foi realizada com melhor taxa para o modelo SW quando realizado a fusão dos pH 3 com pH 7, resultando num modelo com taxa de acerto superior a 90%. Por outro lado, os resultados do MCR-ALS apresentaram ótima recuperação dos compostos fluorescentes presente nos vinhos tinto analisados, e o DD-SIMCA alta capacidade de reconhecimento geográfico. Isso mostra que a metodologia proposta pode ser utilizada como uma ferramenta eficaz para identificação e classificação de amostras de vinhos visando a rastreabilidade desse produto, quando consideradas as mais representativas regiões produtoras de vinho do Estado do Rio Grande do Sul.

**Palavras-chave:** Espectrometria de Fluorescência, Vinho Tinto, Origem Geográfica, Quimiometria.





## ABSTRACT

New methodologies for spectrofluorimetry coupled with chemometric tools were developed, aiming at the classification of red wine samples produced in two regions of Rio Grande do Sul, considering their geographical origin. As the fluorescent properties of a chemical compound can vary depending on the pH, and the structural changes resulting from the pH induce significant changes in its fluorescence spectra, it was possible to separate and identify different chemical compounds as well. In this work, 53 samples of red wine from the Serra Gaúcha region and 20 from the Campanha region were analyzed, covering 10 grape varieties. The registered fluorescence signal corresponds to nine emission matrices (51 variables) of excitation (12 variables) (EEM) registered at different pH (3 to 11) thus generating the data matrix representing the higher order data. These were treated by Multivariate Curve Resolution with Alternating Least Squares (MCR-ALS) and one class method Data Driven Soft Independent Modelling of Class Analogy (DD-SIMCA) to build the classification models. Two pHs (3 and 7) were also selected, as well as the fusion of their data, to represent 1st order data, which were explored through the application of the algorithms (ACO, GA and SW) to select variables, to assist in the recognition of wines using linear discriminant analysis. As a result, it was observed that the prediction was performed with a better rate for the SW model when the fusion of pH 3 with pH 7 was performed, resulting in a model with a hit rate greater than 90%. On the other hand, the results of the MCR-ALS showed an excellent recovery of the fluorescent compounds present in the analyzed red wines, and the DD-SIMCA high capacity for geographic recognition. This shows that the proposed methodology can be used as an effective tool for the identification and classification of wine samples aiming at the traceability of this product, when considered the most representative wine producing regions of the State of Rio Grande do Sul.

**Keywords:** Fluorescence Spectrometry, Red Wine, Geographic Origin, Chemometrics.



## LISTA DE FIGURAS

<b>Figura 1:</b> Esquema do funcionamento do espectrofluorímetro (Adaptado: SKOOG, HOLLER & NIEMAN, 2005).....	35
<b>Figura 2:</b> Representação de objetos diferentes que podem ser construídos com dados medidos multivariados para uma única amostra (Adaptado: OLIVIERI & ESCANDAR, 2014).....	37
<b>Figura 3:</b> Categorização dos dados analíticos (Adaptado: ESCANDAR & OLIVIERI, 2014).....	39
<b>Figura 4:</b> Representação do modelo de resolução de curvas multivariada para a matriz $D$ constituída de $I$ amostras contendo $F$ componentes de medidas em $J$ comprimentos de onda (Adaptado: FILHO, 2011). ....	42
<b>Figura 5:</b> Representação das etapas do MCR-ALS (Adaptado: TAULER, 2005).....	44
<b>Figura 6:</b> Diagrama mostrando a desenvolvimento do tempo que as formigas levam no processo de busca do alimento e o seu o retorno para o ninho do levando o alimento. O ninho e a fonte de comida são otimizados por meio do trabalho conjunto da colônia (Adaptado: GOSS <i>et al.</i> , 1989).....	46
<b>Figura 7:</b> Fluxograma representando as principais etapas do algoritmo genético básico (Adaptado: FILHO & POPPI, 1999). ....	50
<b>Figura 8:</b> Representação de uma codificação espectral na forma de um cromossomo. Neste caso, cada valor de comprimento corresponde a um gene do cromossomo (Adaptado: FILHO & POPPI, 1999). ....	53
<b>Figura 9:</b> Representação da elaboração de vinho registrada pelos egípcios através de pinturas no ano de 1100 a.C. (Adaptado: PRATES, 2018). ....	57
<b>Figura 10:</b> Gráfico da produção total mundial de acordo com os tipos de produtos (Adaptado: OIV, 2017).....	58
<b>Figura 11:</b> Gráficos dos cinco países que consomem metade do vinho do mundo (OIV, 2017).....	59
<b>Figura 12:</b> Esquema de transformação dos dados espectrais realizado.....	75
<b>Figura 13:</b> Gráficos das médias dos sinais espectrais das amostras de treinamento de vinhos tintos, para o pH3, sendo a Classe 1 referente as amostras da Campanha e Classe 2 referente as amostras da Serra Gaúcha. ....	76
<b>Figura 14:</b> Gráfico das variáveis selecionadas para o modelo ACO4 e perfil médio dos espectros das amostras do pH3. ....	79
<b>Figura 15:</b> Gráfico das variáveis selecionadas para o modelo GA5 e perfil médio dos espectros das amostras do pH3. ....	79
<b>Figura 16:</b> Gráfico dos resultados da LDA/ACO4 para as amostras do vinho tinto com pH3 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).....	78
<b>Figura 17:</b> Gráfico dos resultados da LDA/GA5 para as amostras do vinho tinto com pH3 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis	

(treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).....	780
<b>Figura 18:</b> Gráfico das variáveis selecionadas para o modelo SW3 e perfil médio dos espectros das amostras do pH3.....	82
<b>Figura 19:</b> Gráfico dos resultados da LDA/SW3 para as amostras do vinho tinto com pH3 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelho (teste). .....	83
<b>Figura 20:</b> Gráficos das médias dos sinais espectrais das amostras de treinamento de vinhos tintos, para o pH7, sendo a Classe 1 referente as amostras da Campanha e Classe 2 referente as amostras da Serra Gaúcha. ....	84
<b>Figura 21:</b> Gráfico das variáveis selecionadas para o modelo ACO5 e perfil médio dos espectros das amostras do pH7.....	875
<b>Figura 22:</b> Gráfico das variáveis selecionadas para o modelo GA4 e perfil médio dos espectros das amostras do pH7.....	875
<b>Figura 23:</b> Gráfico dos resultados da LDA/ACO5 para as amostras do vinho tinto com pH7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).....	88
<b>Figura 24:</b> Gráfico dos resultados da LDA/GA4 para as amostras do vinho tinto com pH7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).....	88
<b>Figura 25:</b> Gráfico das variáveis selecionadas para o modelo SW4 e perfil médio dos espectros das amostras do pH7.....	88
<b>Figura 26:</b> Gráfico dos resultados da LDA/SW4 para as amostras do vinho tinto com pH7 das duas regiões, sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).....	89
<b>Figura 27:</b> Gráficos das médias dos sinais espectrais das amostras de treinamento de vinhos tintos, para a fusão dos dados do pH3 e do pH7, sendo a Classe 1 referente as amostras da Campanha e Classe 2 referente as amostras da Serra Gaúcha.....	92
<b>Figura 28:</b> Gráfico das variáveis selecionadas para o modelo ACO4 e perfil médio dos espectros das amostras para a fusão dos dados do pH3 e do pH7. ....	94
<b>Figura 29:</b> Gráfico das variáveis selecionadas para o modelo GA2 e perfil médio dos espectros das amostras para a fusão dos dados do pH3 e do pH7.....	95
<b>Figura 30:</b> Gráfico dos resultados da LDA/ACO4 para as amostras do vinho tinto com pH3 e pH7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste). ....	95
<b>Figura 31:</b> Gráfico dos resultados da LDA/GA2 para as amostras do vinho tinto com pH3 e pH 7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).....	96

**Figura 32:** Gráfico das variáveis selecionadas para o modelo SW3 e perfil médio dos espectros das amostras para a fusão dos dados dos do pH3 e do pH7. .... 97

**Figura 33:** Gráfico dos resultados da LDA/SW3 para as amostras do vinho tinto com pH3 e pH 7 das duas regiões, sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste)..... 98



## LISTA DE TABELAS

<b>Tabela 1:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de treinamento para o pH3. ...	77
<b>Tabela 2:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de teste para o pH3. ....	77
<b>Tabela 3:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de treinamento para o pH3. ....	78
<b>Tabela 4:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de teste para o pH3. ....	78
<b>Tabela 5:</b> Resultados da aplicação da formulação <i>Stepwise</i> (SW) para os dados espectrais das amostras de treinamento para o pH3. ....	81
<b>Tabela 6:</b> Resultados da aplicação da formulação <i>Stepwise</i> (SW) para os dados espectrais das amostras de teste para o pH3. ....	82
<b>Tabela 7:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de treinamento para o pH7. ...	85
<b>Tabela 8:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de teste para o pH7. ....	85
<b>Tabela 9:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de treinamento para o pH7. ....	86
<b>Tabela 10:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de teste para o pH7. ....	86
<b>Tabela 11:</b> Resultados da aplicação da formulação <i>Stepwise</i> (SW) para os dados espectrais das amostras de treinamento para o pH7. ....	87
<b>Tabela 12:</b> Resultados da aplicação da formulação <i>Stepwise</i> (SW) para os dados espectrais das amostras de teste para o pH7. ....	88
<b>Tabela 13:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de treinamento para a fusão dos dados do pH3 e do pH7. ....	92
<b>Tabela 14:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de teste para a fusão dos dados do pH3 e do pH7. ....	93
<b>Tabela 15:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de treinamento para a fusão dos dados do pH3 e do pH7. ....	93
<b>Tabela 16:</b> Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de teste para a fusão dos dados do pH3 e do pH7. ....	94
<b>Tabela 17:</b> Resultados da aplicação da formulação <i>Stepwise</i> (SW) para os dados espectrais das amostras de treinamento para a fusão dos dados dos pH3 e pH7. ....	96
<b>Tabela 18:</b> Resultados da aplicação da formulação <i>Stepwise</i> (SW) para os dados espectrais das amostras de teste para a fusão dos dados dos pH3 e pH7. ....	97





## LISTA DE ABREVIATURAS

1. **ACO** Algoritmo Colônia de Formigas (do inglês *Ant Colony Optimization*)
2. **CLAE** Cromatografia Líquida de Alta Eficiência
3. **DAD** Detector de Arranjos de Diodos
4. **EEM** Matrizes de Excitação e Emissão (do inglês *Excitation and Emission Matrices*)
5. **GA** Algoritmo Genético (do inglês *genetic algorithm*)
6. **GCMS/MS** Cromatografia Gasosa Acoplada a Espectrômetro de Massas em Modo *Tandem*
7. **IBGE** Instituto Brasileiro de Geografia e Estatística
8. **IUPAC** União Internacional de Química Pura e Aplicada (do inglês *International Union of Pure and Applied Chemistry*)
9. **LS-SVM** Máquinas de Vetor de Suporte com Mínimos Quadrados (do inglês *Least Squares Support Vector Machine*)
10. **LRM** Limites Máximos de Resíduos
11. **MCR** Resolução de Curvas Multivariadas (do inglês *Multivariate Curve Resolution*)
12. **MCR-ALS** Resolução de Curvas Multivariadas por Mínimos Quadrados Alternantes (do inglês *Multivariate Curve Resolution with Alternating Least Squares*)
13. **OIV** Organização Internacional de Vinha e Vinhos
14. **PARAFAC** Análise de Fatores Paralelos (do inglês *Parallel Factor Analysis*)
15. **PCR** Regressão por Componentes Principais (do inglês *Principal Component regression*)
16. **PCA** Análise por Componentes Principais (do inglês *Principal Component Anaysis*)
17. **PLS** Mínimos Quadrados Parciais (do inglês *Partial Least Squares*)
18. **PNCR** Plano Nacional de Controle de Resíduos e Contaminantes
19. **RFA** Análise de resolução de fatores (do inglês *Resolving Factor Analysis*)
20. **RNA** Rede Neural Artificial
21. **SW** Algoritmo *Stepwise* (ou Formulação *Stepwise*)



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	19
<b>2</b>	<b>OBJETIVOS</b> .....	25
2.1	Objetivo Geral .....	27
2.2	Objetivos Específicos .....	27
<b>3</b>	<b>REVISÃO BIBLIOGRÁFICA</b> .....	29
3.1	Quimiometria e Técnica Analítica .....	29
3.1.1	Quimiometria.....	29
3.1.2	Controle de Qualidade.....	29
3.1.3	Espectroscopia de Fluorescência.....	34
3.2	Análise Multivariada Multivias .....	36
3.2.1	Calibração em Química Analítica .....	36
3.2.2	Classificação dos Dados Analíticos .....	38
3.2.3	Resolução Multivariada de Curvas (MCR).....	39
3.2.4	Resolução Multivariada de Curvas por Mínimos Quadrados Alternantes (MCR-ALS).....	42
3.2.5	Algoritmos de Seleção de Variáveis .....	45
3.2.5.1	Algoritmo Colônias de Formigas (ACO).....	45
3.2.5.2	Algoritmo Genético (GA).....	47
3.2.5.3	Algoritmo <i>Stepwise</i> (SW) .....	54
3.3	Matriz Vinho Tinto.....	56
3.3.1	Vinho Tinto .....	56
3.3.2	História de Uvas e Vinhos.....	56
3.3.3	Vitivinicultura Dados Macroeconômicos.....	59
3.3.4	História da Vitivinicultura no Brasil .....	59
3.3.5	Vitivinicultura no Estado do Rio Grande do Sul.....	65
3.3.5.1	Vitivinicultura na Serra Gaúcha.....	65
3.3.5.2	Vitivinicultura na Região da Campanha .....	66
3.3.6	Composição e pH do Vinho .....	67
3.3.7	Vinho e Saúde .....	67
<b>4</b>	<b>CAPÍTULO – Modelos de Seleção de Variáveis</b> .....	69

4.1	Apresentação.....	73
4.1.1	Introdução.....	73
4.1.2	Materiais e Métodos .....	74
4.1.2.1	Solução Tampão e Produtos Químicos .....	74
4.1.2.2	Amostras de Vinhos .....	74
4.1.2.3	Método Analítico .....	74
4.1.2.4	Processos Quimiométricos.....	75
4.1.3	Resultados e Discussão .....	76
4.1.3.1	Resultados das Aplicações dos Algoritmos .....	76
4.1.3.1.1.	Resultados da Seleção do pH3.....	74
4.1.3.1.2.	Resultados da Seleção do pH7.....	83
4.1.3.1.3.	Resultados da Seleção para a fusão dos pH3 e pH7 .....	839
<b>5</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>99</b>
<b>6</b>	<b>CAPÍTULO - ARTIGO CIENTÍFICO – Southern Brazilian Red Wine Geographical Origin by Means of EEM-pH Four Way Data Modelling Coupled With One Class Approach .....</b>	<b>117</b>
6.1	Apresentação .....	119
6.2	Artigo.....	119
<b>7</b>	<b>DISCUSSÃO GERAL .....</b>	<b>145</b>
<b>8</b>	<b>CONCLUSÃO.....</b>	<b>151</b>
<b>9</b>	<b>ANEXOS.....</b>	<b>155</b>

---

## **1 INTRODUÇÃO**

---



A espectroscopia de fluorescência vem sendo utilizada cada vez mais em diversas áreas da ciência como um método alternativo para análise de compostos que possuem esta propriedade de fluorescência. Com a combinação de técnicas espectroscópicas e quimiométricas, este método apresenta vantagens várias em relação as outras abordagens espectroscópicas como rapidez, baixo custo, em alguns casos não precisa da etapa de preparação da amostra, baixo consumo de reagentes e potencial de portabilidade (RODRÍGUEZ *et al.*, 2011; AZCARATE *et al.*, 2015; GREDILLA *et al.*, 2016; SAAD *et al.*, 2016).

A espectroscopia de fluorescência pode ser utilizada para análise de diferentes tipos de matrizes. No caso em particular do vinho, destacam-se algumas vantagens como alta sensibilidade e seletividade, essas merecendo atenção em comparação com outras abordagens espectroscópicas (LAKOWICZ, 2006). As propriedades fluorescentes de um composto químico podem variar em função do pH e, assim, o registro do sinal de fluorescência em diferentes pH permite uma caracterização completa dos compostos fluorescentes presentes no vinho (GREDILLA *et al.*, 2016; LAKOWICZ, 2006; SAAD *et al.*, 2016).

Também é importante ressaltar que a versatilidade dos espectrofluorímetros permite o registro do sinal na forma de um valor de fluorescência simples, sendo obtido para comprimentos de onda de excitação e emissão específicos (espectros de emissão cuja amostra foi excitada em um comprimento de onda de excitação fixo) e mapa de fluorescência (espectros obtidos em comprimentos de onda de excitação de distanciamento). Estes tipos de dados são dispostos em forma de matrizes de emissão e excitação (EEM). Dessa forma, os dados são adequados para aplicação de técnicas quimiométricas que podem ser utilizados no controle de qualidade de vinhos tintos brasileiros visando a rastreabilidade de sua origem geográfica (LAKOWICZ, 2006; KUMAR & MISHRA, 2012).

Por outro lado, sabe-se que o valor comercial dos vinhos é geralmente baseado na sua origem geográfica e na variedade de uva. Algumas regiões são reconhecidas por produzir excelentes vinhos com alto valor comercial. A origem dos produtos é questão importante para a indústria alimentar, não só para os consumidores, mas também para os produtores e distribuidores (LUYKX & RUTH, 2008). Devido ao seu grande consumo, o vinho torna-se alvo de investigação devido a fraudes, incluindo adulteração, declaração de idade falsa e dados de origem geográfica. Assim, muitos esforços vêm sendo feito para

## INTRODUÇÃO

---

desenvolver métodos que consigam acessar a qualidade do vinho e evitar fraudes (ÁLVAREZ *et al.*, 2007).

Neste contexto, o presente trabalho tem por finalidade o desenvolvimento de metodologias analíticas visando identificar os compostos e classificar os vinhos tintos varietais abordando duas principais regiões produtoras de vinho do Rio Grande do Sul e através da aplicação de distintas ferramentas quimiométricas aos dados de fluorimetria. No Brasil, ainda não existem trabalhos que identifiquem compostos químicos nessas matrizes aplicando essa técnica.



---

## **2 OBJETIVOS**

---



## 2.1 Objetivo Geral

- Desenvolver metodologias analíticas para identificação da origem geográfica de vinhos tintos produzidos no Rio Grande do Sul empregando dados de espectrometria por fluorescência e ferramentas quimiométricas.

## 2.2 Objetivos Específicos

- Desenvolver modelos de classificação geográfica a partir dos dados de fluorescência dos compostos presentes nos vinhos (Serra Gaúcha e Campanha) empregando análise linear discriminante (LDA);
- Comparar o desempenho dos métodos de seleção de variáveis como o algoritmo colônia de formigas (ACO), algoritmo genético (GA) e formulação *stepwise* (SW) visando a classificação de origem geográfica de vinhos tintos;
- Determinar os perfis dos compostos fluorescentes presentes nos vinhos tintos (Serra Gaúcha e outras regiões) em diferentes pHs empregando matrizes de fluorescência e resolução multivariada de curvas por mínimos quadrados alternantes (MCR-ALS);
- Desenvolver modelo *one class* visando a certificação de origem geográfica (Serra Gaúcha), a partir de dados de fluorescência dos vinhos tintos varietais, empregando o método *data driven soft independent modelling of class analogy* (DD-SIMCA).



---

### **3 REVISÃO BIBLIOGRÁFICA**

---



### 3.1 Quimiometria e Técnica Analítica

A espectroscopia de fluorescência molecular foi a técnica analítica escolhida para desenvolver esse trabalho; junto a essa técnica foram usadas as ferramentas quimiométricas para tratamentos e análises dos dados no controle de qualidade de vinho tinto brasileiro produzido no Rio Grande do Sul. Neste tópico são apresentados os fundamentos sobre a quimiometria e sua aplicação no controle de qualidade para a matriz estudada. Além disso, são abordados os princípios da espectroscopia de fluorescência e seu funcionamento.

#### 3.1.1 Quimiometria

A quimiometria é uma ferramenta criada pelo químico Svante Wold, em 1974, usada para descrever as técnicas e operações associadas ao tratamento matemático e a interpretação dos dados químicos. A quimiometria é uma área da química que usa métodos matemáticos, estatísticos e outros que empregam lógica formal com os seguintes objetivos:

- 1) Conceber ou selecionar procedimentos de medição ótimas e experiências.
- 2) Fornecer a quantidade máxima de informações por meio da análise de dados químicas relevantes (FERREIRA *et al.*, 1999; WORKMAN *et al.*, 2011).

Existe uma grande demanda por resultados, rápidos, expressos de forma simples e conclusivos a partir de um grande volume de dados. Isto faz com que a quimiometria esteja presente em praticamente todas as áreas de aplicação da química analítica, podendo extrair o máximo de informação possível dos dados químicos analisados (FERREIRA *et al.*, 1999; WORKMAN *et al.*, 2011).

#### 3.1.2 Controle de Qualidade

A grande demanda de produtos alimentares seguros de alta qualidade fez com que a Europa desenvolvesse leis que regulamentam a autenticidade e origem geográfica dos produtos a fim de assegurar a sua qualidade e proteger os consumidores e produtores de possíveis fraudes (JACKSON, 2008b).

O vinho é uma bebida consumida no mundo todo e está entre os produtos alimentares mais estudados e protegidos com uma boa regulamentação, certificação de origem e procedência. Por isto, a rastreabilidade geográfica do vinho ganhou importância na sociedade, originando a necessidade de protocolos científicos válidos capazes de traçar a proveniência geográfica dos vinhos (JACKSON, 2008b).

O valor comercial do vinho está normalmente relacionado à sua origem geográfica e variedade da uva. Algumas regiões são reconhecidas por produzirem excelentes vinhos com alto valor comercial. A utilização falsa de indicações geográficas por partes não autorizadas é prejudicial para os consumidores e produtores legítimos (LUYKX & VAN, 2008).

Estudos realizados nos anos de 1990-2000 apontam que os consumidores recuperaram interesse em alimentos fortemente identificados com um local de origem (FSA, 2001), e por exemplo, uma avaliação suíça mostrou que a origem dos alimentos é um fator importante na decisão de compra de 82% dos clientes (FRANKE *et al.*, 2005).

Existem estudos fundamentados na impressão digital química que o *terroir* deixa na composição do produto final, permitindo que o cientista defina uma correlação entre o alimento e a sua área de produção. Sabe-se que uma rigorosa correlação entre o vinho e o seu *terroir* que é um conjunto de fatores no espaço onde ele se desenvolve que inclui *terroir*, específicas (uma área precisa onde o clima, pedologia, drenagem, biodiversidade, práticas vitivinícolas que são aplicadas, os fatores geológicos e vitícolas) que interagem e dão ao vinho caráter único da sua tipicidade (VAUDOUR, 2002). O *terroir* faz com que o produto tenha um diferencial, agregando valor à devido a sua identidade territorial (FLORES & MEDEIROS, 2013).

Existem diversas abordagens na literatura sobre o desenvolvimento de novas técnicas, cada vez mais sofisticadas, para determinar a origem geográfica dos produtos agrícolas, sendo tal fato de grande importância para os consumidores, agricultores, varejistas e autoridades administrativas (LUYKX & VAN, 2008).

Relatórios sobre os métodos analíticos para determinar a origem geográfica de produtos agrícolas têm aumentado desde os anos 80. A quimiometria associada a técnicas analíticas vem sendo muito utilizada na área de alimentos, principalmente para controle de qualidade. Os dados fornecidos por instrumentos analíticos têm capacidade de determinar mais de um componente por vez em uma amostra e, assim, podem estabelecer ligações sobre a origem alimentar. Se os componentes têm poder discriminatório



suficiente, o conjunto de suas concentrações formará um padrão característico ou "impressão digital" relativas à origem geográfica da amostra (LUYKX & VAN, 2008).

Existem diferentes trabalhos que empregam múltiplas técnicas analíticas utilizando ferramentas quimiométricas para análises de diversos alimentos, como por exemplo, vinagres de vinho (RÍOS-REINA *et al.*, 2017), doces (CEBI *et al.*, 2018), suplementos alimentares (WALKOWIAK *et al.*, 2019), vinhos (SAAVEDRA, 2011), hambúrgueres de frango (FERNANDES *et al.*, 2018), gengibres (SHUKLA *et al.*, 2019), batatas doces (SANTOS, 2017), grãos de arroz (RUNGE *et al.*, 2019), óleos (LIU *et al.*, 2019), óleos e frutos, (ELSAYED *et al.*, 2019) entre outros.

No caso do vinho, encontram-se várias abordagens analíticas na literatura para estabelecer a composição e autenticidade do vinho de acordo com sua origem geográfica e a variedade de uva, por exemplo; a cromatografia líquida e gasosa acopladas a vários detectores (massa -MS, arranjos de diodos -DAD, MS-MS e fluorescência) (VILLIERS *et al.*, 2004; RODRÍGUEZ-DÍAS, AGUILLAR-CABALLOS & GÓMEZ-HENS, 2006; TREDoux *et al.*, 2008; BALLABIO, *et al.*, 2008; ALVES, SAURINA, 2010; CUADROS-INOSTROZA, 2010; ALVES, NASCIMENTO & NOGUERIA, 2011; VACLAVIK *et al.*, 2011; DIMITROVSKA, TOMOVSKA & BOCEVSKA, 2013), espectrometria atômica, espectrometria de massas com plasma indutivamente acoplado (ICP-MS), emissão óptica (ICP-OES) (JURADO, 2012; SHEN *et al.*, 2013; ŠELIH, ŠALA & DRGAN, 2014).

Estas técnicas analíticas são abordagens mais convencionais e eficazes na elucidação da composição orgânica e elementar dos vinhos, mas necessitam de equipamentos, custo de preparação de amostras e alta manutenção, o que pode representar um fator limitante para pequenas indústrias ou produtores (JURADO, 2012; SHEN *et al.*, 2013; ŠELIH, ŠALA & DRGAN, 2014).

A espectroscopia de fluorescência em conjunto com a quimiometria pode ser uma alternativa de metodologia para esse tipo de análise, apresentando diversas vantagens como rapidez, preparo simples de amostras e baixo custo quando comparada às atualmente empregadas (SKOOG, HOLLER & NIEMAN, 2002; GÓMEZ & CALLAO, 2008; HORVAT *et al.*, 2012).

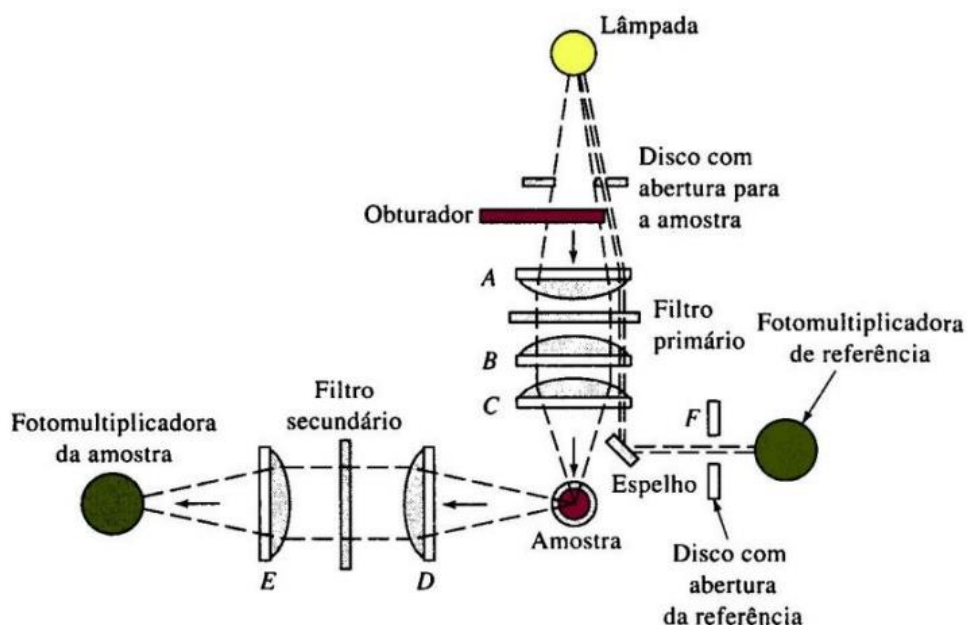
### 3.1.3 Espectroscopia de Fluorescência

A espectroscopia de fluorescência é uma técnica amplamente utilizada em muitos campos relevantes para as ciências químicas, biológicas e médicas (CHRISTIAN *et al.*, 1981; DESTRAMPE & HIEFTJE, 1993; SOPER *et al.*, 1994).

A espectroscopia de fluorescência, também chamada de espectrofluorimetria ou simplesmente fluorimetria, é uma técnica analítica empregada para análise quantitativa e qualitativa para analitos que são capazes de emitir de fluorescência (SKOOG, HOLLER & NIEMAN, 2002; VOGEL *et al.*, 2002).

A fluorescência é um fenômeno de absorção de energia radiante e da posterior emissão de parte da energia. Existem várias vantagens dessa técnica, por exemplo, a elevada sensibilidade e a maior seletividade, já que nem todas as substâncias que absorvem radiação eletromagnética são capazes de emitir fluorescência (SKOOG, HOLLER & NIEMAN, 2002). A fluorescência é um processo rápido que se completa em cerca de  $10^{-5}$  a  $10^{-10}$  s após a excitação (SKOOG, HOLLER & CROUCH, 2009; SANTOS & GIL, 2010).

A atividade fluorescente do analito ocorre quando ele absorve a radiação em um determinado comprimento de onda  $\lambda$ , ou seja, no comprimento de onda de excitação ( $\lambda_{ex}$ ) e ao retornarem para o seu estado fundamental, a energia residual é emitida na forma de fóton. O comprimento de onda de excitação pode ser diferenciado do comprimento de emissão ( $\lambda_{em}$ ), pois o fóton emitido apresenta menor quantidade de energia que os fótons da excitação (absorção). Cada molécula fluorescente possui um comprimento de onda característico de excitação e emissão, porém o comprimento de emissão é maior que o de excitação devido à perda de energia (SKOOG, HOLLER & NIEMAN, 2002; LAKOWICZ, 2009).



**Figura 1:** Esquema do funcionamento do espectrofluorímetro (Adaptado: SKOOG, HOLLER & NIEMAN, 2005).

O espectro de absorção de um cromóforo é, sobretudo, determinado pela estrutura química da molécula. Alguns fatores ambientais podem influenciar nas mudanças detectáveis no comprimento de onda máximo e na emissão. O pH do solvente, que determina o estado de ionização dos cromóforos ionizáveis, pode ter influência. Geralmente, a fluorescência de um composto aromático com substituintes ácidos ou básicos é dependente do pH. Assim, o comprimento de onda, bem como a intensidade de emissão, será diferente para as formas protonadas e desprotonadas (SKOOG, HOLLER & CROUCH, 2009).

Em compostos contendo grupos aromáticos funcionais com baixa energia é encontrada fluorescência mais intensa. Os compostos contendo estruturas alifáticas e carbonilas alicíclicas ou estruturas de ligações duplas altamente conjugadas também podem apresentar fluorescência, mas em menor número que nos sistemas aromáticos (SKOOG, HOLLER & CROUCH, 2009).

### 3.2 Análise Multivariada Multivias

Nesta seção é apresentada uma revisão de calibração em química analítica e das classificações dos dados analíticos de ordem superior. Além disso, são descritos os princípios fundamentais dos algoritmos MCR-ALS, ACO, GA e SW empregado neste trabalho.

#### 3.2.1 Calibração em Química Analítica

De acordo com a União Internacional de Química Pura e Aplicada (IUPAC), a calibração é, em um sentido geral, “uma operação que relaciona uma quantidade de saída com uma quantidade de entrada para um sistema de medição sob certas condições” (IUPAC, 1998).

As grandezas de entrada de nosso interesse primário, ou seja, em calibração analítica, são as concentrações de um constituinte da amostra de interesse (o analito), enquanto as quantidades de sinais analíticos ou respostas entregues por instrumentos analíticos será os sinais instrumentais. Assim, a calibração significará a operação de relacionar sinais instrumentais para analisar as concentrações (OLIVIERI & ESCANDAR, 2014).

Existem dois tipos de calibração: univariada e multivariada. Na calibração univariada, uma única resposta instrumental é relacionada à propriedade de interesse. Em química analítica, a calibração univariada emprega uma curva de calibração como método geral para a determinação de concentração de um constituinte em uma amostra desconhecida (IUPAC, 1998). Na calibração multivariada, um conjunto de respostas instrumentais é relacionado com uma ou várias propriedades (pode ser ou não a concentração do analito) das amostras (SENNA *et al.*, 2005).

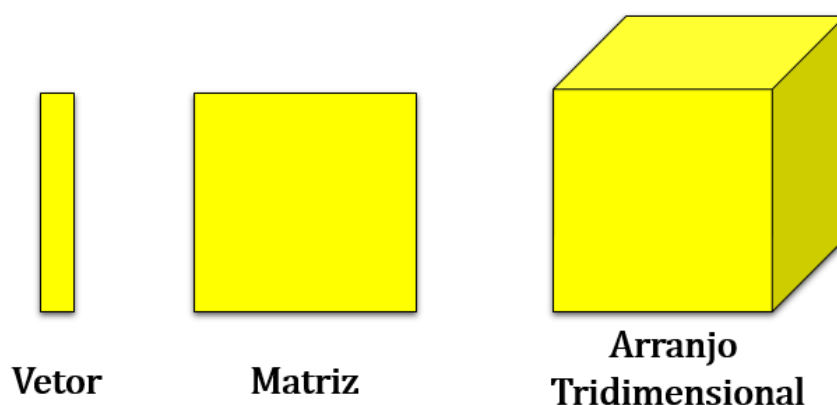
A curva de calibração é um gráfico de resposta instrumental em função da concentração do analito. Normalmente, o analista prepara uma série de padrões, ao longo de uma série de concentrações próximas da concentração esperada do analito, em uma amostra desconhecida. As concentrações dos padrões estão dentro da faixa de trabalho empregado na técnica analítica. Analisar cada um destes padrões, usando a técnica escolhida, produzirá uma série de medições (OLIVIERI & ESCANDAR, 2014).

Para a maioria das análises, um gráfico de resposta instrumental versus concentração de analito mostrará uma relação linear dentro de uma determinada faixa de concentração. Depois de medir a resposta para a amostra desconhecida, a interpolação da curva de calibração permite encontrar a concentração do analito (OLIVIERI & ESCANDAR, 2014).

Um processo de calibração mais geral envolve a relação entre as concentrações de vários constituintes, numa amostra de teste, e múltiplas respostas medidas, isto é, multivariadas. Em contraste com a calibração univariada, que trabalha com uma única resposta instrumental medida em cada amostra experimental, a calibração multivariada funciona com muitos sinais diferentes para cada amostra (SENNA *et al.*, 2005; OLIVIERI & ESCANDAR, 2014).

Dependendo da configuração instrumental, os dados entregues para uma única amostra podem ter graus de complexidade diferentes. Os dados multivariados mais simples são aqueles produzidos em forma vetorial, ou seja, como uma série de respostas, que podem ser colocadas uma em cima da outra para gerar um objeto matemático conhecido como vetor coluna. Este objeto também é referido como tendo um único “modo” ou “direção” (OLIVIERI & ESCANDAR, 2014).

Matrizes de dados mais complexas que vetores podem ocorrer. A crescente complexidade deriva de configurações instrumentais cada vez mais complexas, que geram dados que podem ser organizados em objetos com modos adicionais, como uma matriz ou um arranjo tridimensional (**Figura 2**) (OLIVIERI & ESCANDAR, 2014).



**Figura 2:** Representação de objetos diferentes que podem ser construídos com dados medidos multivariados para uma única amostra (Adaptado: OLIVIERI & ESCANDAR, 2014).

Dentro da calibração multivariada existe a calibração multivias, que emprega dados obtidos em pelo menos dois arranjos de sensores. Na técnica de espectroscopia de fluorescência por exemplo, podem ser usadas as medidas de fluorescência para gerar uma matriz de excitação e emissão (EEM) formando um arranjo de dados tridimensionais para um conjunto de amostras (MURPHY *et al.*, 2013).

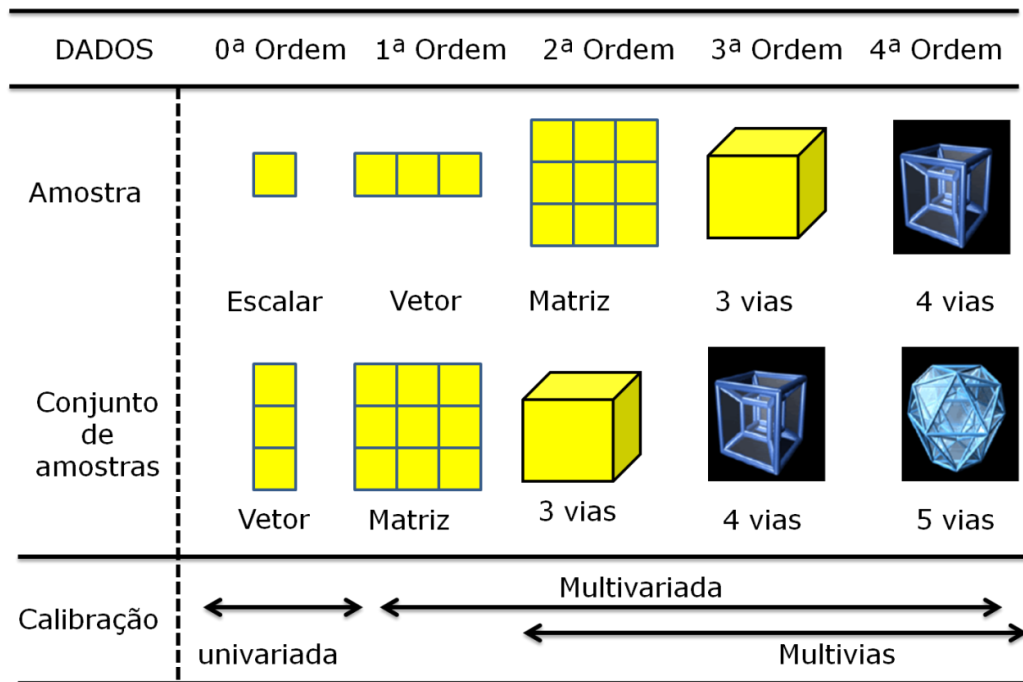
Esse tipo de matriz apresenta dados muito complexos, por isso é aplicado a calibração multivias, que tem como vantagens a capacidade de técnica determinar mais de um analito no mesmo ensaio e permitir a quantificação do analito na presença de interferentes (OLIVIERI & ESCANDAR, 2014).

Na literatura, existem vários algoritmos que são empregados na calibração de multivias. Os mais usados para dados de calibração multivias (dois ou mais arranjos) são: Análise de Fatores Paralelos (PARAFAC) e Resolução de Curva Multivariada com Mínimos Quadrados Alternantes (MCR-ALS) (BRO, 1997; JUAN, RUTAN & TAULER, 2009).

### 3.2.2 Classificação dos Dados Analíticos

Em química analítica, os métodos de calibração usados podem ser classificados de acordo com a dimensão dos dados analisados e o processo de calibração poder ser classificado como univariado, multivariado e multivias (BOOKSH & KOWALSKI, 1994).

Os métodos de ordem zero (vetor **Figura 3**) são usados para tratar dados univariados gerados por instrumentos tais como, eletrodos íon-seletivos, pHmetros e colorimétricos. Nestes casos, a resposta medida para cada amostra é um valor escalar (tensor de ordem zero). Estes métodos não fornecem resultados aceitáveis na presença de interferentes, pois exigem total seletividade para o analito de interesse (SENNA *et al.*, 2005).



**Figura 3:** Categorização dos dados analíticos (Adaptado: ESCANDAR & OLIVIERI, 2014).

Os métodos de calibração de primeira ordem (**Figura 3**) podem ser usados para tratar dados multivariados gerados por instrumentos tais como, espectrômetros e cromatógrafos, cujas respostas fornecem um vetor (tensor de primeira ordem) de dados para cada amostra. Nestes casos, é possível a calibração na presença de interferentes, desde que estes estejam presentes no conjunto de calibração usado para construir o modelo (SENNA *et al.*, 2005). Dados de ordem igual ou superior a dois são chamados de dados de ordem superior (NIKOLAJSSEN *et al.*, 2003; OLIVIERI *et al.*, 2004).

Já os dados de primeira ordem são processados por procedimentos de calibração multivariada de primeira ordem aplicando algoritmos lineares, como a regressão por componentes principais (PCR), mínimos quadrados parciais (PLS) (MARTENS & NAES, 1989), ou algoritmos não lineares, como rede neural artificial (RNA) (DESPAGNE & MASSART, 1998) e mínimos quadrados parciais por máquina de vetores de suporte (LS-SVM) (SUYKENS & VANDEWALLE, 1998).

Instrumentos que fornecem como resposta uma matriz (tensor de segunda ordem) de dados para cada amostra geram dados de segunda ordem (arranjo tridimensional **Figura 3**). Instrumentos, tais como, cromatografia gasosa acoplada à espectrometria de massas (GCMS-MS), cromatografia líquida de alta eficiência com detecção por arranjos de

diodos (CLAE-DAD), espectrofluorimetria (emissão/excitação), análise de injeção em fluxo com detecção espectrofotométria ou análise de imagens, geram uma grande quantidade de informações analisadas (BOOKSH & KOWALSKI, 1994; SENNA *et al.*, 2005).

Os métodos de calibração de segunda ordem, aplicados a esses tipos de dados, apresentam uma série de vantagens sobre os demais métodos que são chamadas “vantagem de segunda ordem”. A principal vantagem é a possibilidade de calibração, sob certas condições, na presença de interferentes desconhecidos, que não estejam presentes no conjunto de calibração. Outra vantagem é a possibilidade de obtenção dos espectros puros de cada componente do sistema (ou outros sinais, dependendo da técnica empregada), aplicando um número mínimo de restrições ao modelo (BOOKSH & KOWALSKI, 1994; SENNA *et al.*, 2005).

Os métodos de ordem igual ou superior a três são possíveis, no entanto, ainda não são encontrados frequentemente na literatura química. Superfícies de fluorescência adquiridas em função do tempo para cada amostra é um exemplo atual de dados de terceira ordem presentes na literatura (NIKOLAJSEN *et al.*, 2003; OLIVIERI *et al.*, 2004).

Dados com mais vias podem ser obtidos por variação de pH, por exemplo, com o uso de técnicas analíticas bidimensionais. Matrizes por fluorescência (excitação-emissão) podem ser registradas com distintos pH produzindo um tensor de dados para cada amostra (WU *et al.*, 2014).

Dados de segunda ordem podem ser tratados por muitos algoritmos. Quando a ordem dos dados aumenta, os dados matemáticos processados tornam-se mais complexos. No entanto, a calibração de segunda ordem geralmente tem uma propriedade especial (vantagem de segunda ordem) em relação à presença de interferentes (BOOKSH & KOWALSKI, 1994).

Na calibração multivias, é importante definir diferentes categorias de amostra e constituintes. As amostras que possuem sinais sobrepostos do analito podem ser considerados como potenciais interferentes (ESCANDAR & OLIVIERI, 2014).

As amostras de calibração são utilizadas para estabelecer a relação entre as concentrações conhecidas do analito e os sinais instrumentais coletados. A validação é empregada para verificar o comportamento da calibração sob condições controladas e os valores de concentração do analito (ESCANDAR & OLIVIERI, 2014).



A composição química das amostras de validação deve ser semelhante às amostras de calibração. Já as amostras de testes podem ter constituintes adicionais aos presentes na calibração e na validação e são utilizados na calibração multivias para avaliar a capacidade que o modelo tem de diferenciar a presença de interferentes (ESCANDAR & OLIVIERI, 2014).

Depois dessa etapa, é recomendável ter um conjunto de amostras, onde a composição é similar às amostras de teste, para as quais as concentrações do analito não seja conhecida, mas podem ser determinadas por uma técnica analítica de referência. Esse conjunto real tem como objetivo avaliar se a precisão e a exatidão do modelo de calibração multivias são semelhantes à técnica de referência. Assim, podem ser comparados os componentes presentes em ambos os modelos de calibração e validação (amostras reais e teste) (ESCANDAR & OLIVIERI, 2014).

### 3.2.3 Resolução Multivariada de Curvas (MCR)

O método quimiométrico de Resolução de Curvas Multivariada (MCR) é uma denominação genérica de uma família de técnicas que tem como função resolver análises de misturas por meio da determinação de contribuições individuais da espécie química (RUTAN, JUAN & TAULER, 2009; MARÇO *et al.*, 2011).

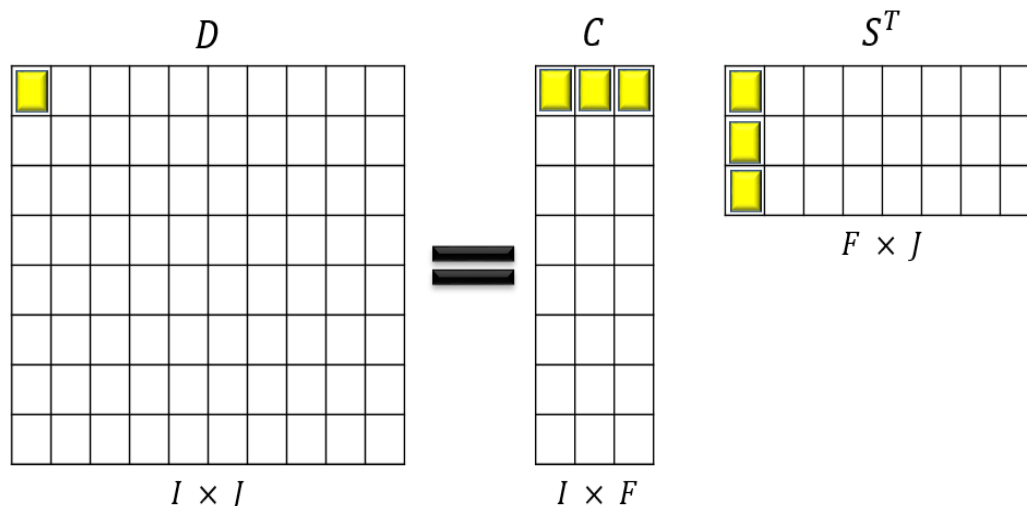
Podemos representar de um modo geral o MCR pela Equação:

$$D = CS^T + E \quad (1)$$

onde  $D$  é a matriz de resposta instrumental,  $C$  representa a matriz de concentração relativa e  $S$  está relacionado a uma matriz de espectros puros. O método MCR realiza a recuperação de dados não seletivos obtidos por um determinado instrumento ( $D$ ) em contribuições reais dos componentes puros do sistema ( $C$  e  $S^T$ ) e mais a informação não modelada, que é o ruído instrumental ( $E$ ) (RUTAN, JUAN & TAULER, 2009; MARÇO *et al.*, 2011).

O MCR realiza a decomposição da matriz que utiliza-se de comprimento de onda dos dados experimentais  $D$  ( $I \times J$ ), de dimensão  $I$  amostras por  $J$  comprimentos de onda. O produto das matrizes  $C$  e  $S^T$ , representa para cada elemento da matriz um produto interno entre um vetor e linha na matriz  $C$ , que possui a concentração de cada espécie da

amostra e um vetor e linha na matriz  $S^T$  (**Figura 4**) que está relacionado às absorvidades molares de cada espécie no comprimento de onda apropriada (MARÇO *et al.*, 2011; FILHO, 2011).



**Figura 4:** Representação do modelo de resolução de curvas multivariada para a matriz  $D$  constituída de  $I$  amostras contendo  $F$  componentes de medidas em  $J$  comprimentos de onda (Adaptado: FILHO, 2011).

### 3.2.4 Resolução Multivariada de Curvas por Mínimos Quadrados Alternantes (MCR-ALS)

Para o processamento de dados analíticos, o método de Resolução Multivariada de Curvas por Mínimos Quadrados Alternantes (MCR-ALS) vem sendo muito utilizado. A partir do ano de 1993, o MCR-ALS tem ganhado destaque em dados de ordem superior, principalmente para resolver misturas de sinais analíticos (TAULER & BARCELÒ, 1993; JUAN & TAULER, 2006; PARASTAR & TAULER, 2014; PEÑA *et al.*, 2015).

O método de análise de dados MCR-ALS, pode ser aplicado a uma única matriz (dados de primeira ordem) ou para matrizes de dados para cada amostra (dados de segunda ordem), determinando a recuperação de valores de concentração relativa e os espectros puros dos componentes relacionados as tais concentrações dentro da amostra, a partir da matriz de dados que contém os valores para as variáveis analisadas (JUAN & TAULER, 2006; MARÇO *et al.*, 2014; PARASTAR & TAULER, 2014).

O MCR-ALS é um tipo de MCR cuja resolução iterativa da **Equação (1)** é realizada por ALS de forma restrita, que calcula as matrizes de concentração  $C$  e espectral  $S$  ajustando a matriz de dados  $D$ . Esses cálculos são feitos através de uma série de interações matemáticas que são repetidas em ciclos, até que um critério bem definido seja atingido (WINDIG *et al.*, 2005; CAO, HARRINGTON & LIU, 2005; AZZOUZ & TAULER, 2008; MARÇO *et al.*, 2011).

Mas, para aplicação desse método, antes, devemos observar se os dados experimentais do modelo seguem duas condições. A primeira, se o sinal analítico obedece a uma relação semelhante a lei de Beer-Lambert para as misturas, ou seja, se os dados têm relação linear com a concentração. Para a segunda condição, deve ser verificado o posto da matriz, esse deve ser igual ao número de espécies que produzem o sinal analítico presentes nas misturas. Este número corresponde ao número de linhas e colunas linearmente independentes, ou seja, o número de vetores que não podem ser escritos como uma combinação linear dos outros. Como exemplo de técnica analítica, podemos citar a espectroscopia de fluorescência, que produz repostas linearmente aditivas que podem ser tratadas dessa maneira (RUTAN, JUAN, A. & TAULER, 2009; MARÇO *et al.*, 2011; OLIVIERI & ESCANDAR, 2014).

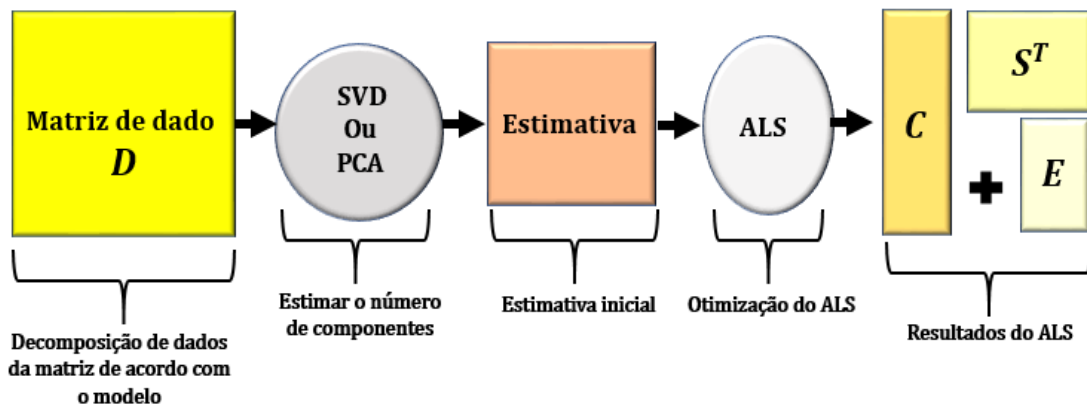
Existem vários métodos numéricos de decomposição de matrizes ou tensores, conforme o modelo, esses métodos são classificados em iterativos e não-iterativos. Os métodos iterativos, como MCR-ALS, são mais populares devido a sua flexibilidade para tratar diversos arranjos de dados e problemas químicos, além de acomodar informações externas no processo de resolução (MARÇO *et al.*, 2011; FILHO, 2011; PEÑA *et al.*, 2015).

Por meio da MCR, o cálculo realizado por ALS garante o melhor ajuste do modelo e esse processo é chamado de otimização. A otimização iterativa, acontece de modo convergente entre a matriz  $D_{est}$  a partir do cálculo de  $C$  e  $S$  e a matriz real  $D$ , tal que o erro  $E$  seja o mínimo. Essa otimização é adaptada conforme as informações químicas e matemáticas no processo sob forma de restrições (JIANG & OZAKI, 2002; PEÑA *et al.*, 2015; YASMIN, 2017).

As restrições são as propriedades que dão forma à linha e perfil das colunas no modelo do conjunto de dados. As restrições são implementadas como condições matemáticas para direcionar o processo de otimização do MCR para as soluções finais. Essas restrições estão ligadas a propriedades das medições do conjunto de dados e outras condições dos modelos. Existe várias restrições, entre elas a não negatividade é uma das

mais usuais, tendo como objetivo fazer com que os perfis sejam formados por valores positivos e pode ser implementada substituindo valores negativos por zeros por exemplo (JUAN, JAUMOT & TAULER, 2014; PEÑA *et al.*, 2015).

Para os dados de primeira ordem, a estimativa pode ser feita por meio de percentagem de variância explicada por Decomposição em Valores Singulares (SVD) (LINDER & SUNDBERG, 1998; MILFORD & SANDELL, 2014) ou por Análise de Componentes Principais (PCA) (WOLD, ESBENSEN & GELADI, 1987) e para os dados de segunda ordem pode ser estimado a partir de Análise de Componentes Principais de Multi-modo (MPCA) (MARÇO *et al.*, 2011).



**Figura 5:** Representação das etapas do MCR-ALS (Adaptado: TAULER, 2005).

Podemos representar o Erro ( $E$ ) pela equação a seguir:

$$E = (D - CS^T)^2 \quad (2)$$

Nos resultados obtidos do MCR é possível observar a qualidade por meio da matriz reconstituída  $D^*$  com a matriz  $D$  original. Existem dois indicadores que podemos analisar essa qualidade a percentagem de falta de ajuste (%LOF) (**Equação 3**) e percentagem de variância explicada ( $R^2$ ) (**Equação 4**) (MARÇO *et al.*, 2011).

$$\%LOF = 100 \times \sqrt{\frac{\sum e_{ij}^2}{\sum d_{ij}^2}} \quad (3)$$

$$R^2 = 100 \times \left( 1 - \frac{\sum e_{ij}^2}{\sum d_{ij}^2} \right) \quad (4)$$

Apesar dos resultados satisfatórios do MCR, alguns problemas de ambiguidade rotacional podem ocorrer na modelagem. Matematicamente, isso acontece porque podem existir várias matrizes  $C$  que ao multiplicar por matriz  $S^T$  resulte na matriz aumentada  $D$  com um resíduo  $E$  baixo e homocedástico (MARÇO *et al.*, 2011; NETO, 2018).

São 3 ambiguidades do MCR-ALS: a de posição em que o perfil de concentração de um analito é relacionado com o perfil espectral de outro constituinte. Para esse problema, podemos resolver facilmente utilizando a organização dos perfis por ordem de variáveis puras. A segunda é de intensidade, que um perfil do analito pode ser mais intenso que o do constituinte, quanto no modo de concentração quanto no espectral. A solução para isso, é normalização de um dos modos, mais usual à normalização dos modos espectrais. Por última e mais difícil de solucionar é ambiguidade rotacional, que faz a recuperação de um perfil que não corresponde a um analito em nenhum dos modos. Para esse problema, são aplicadas as restrições (OLIVIERI, WU & YU, 2009; MARÇO *et al.*, 2011; JAUMONT, JUAN & TAULER, 2015).

### 3.2.5 Algoritmos de Seleção de Variáveis

#### 3.2.5.1 Algoritmo Colônias de Formigas (ACO)

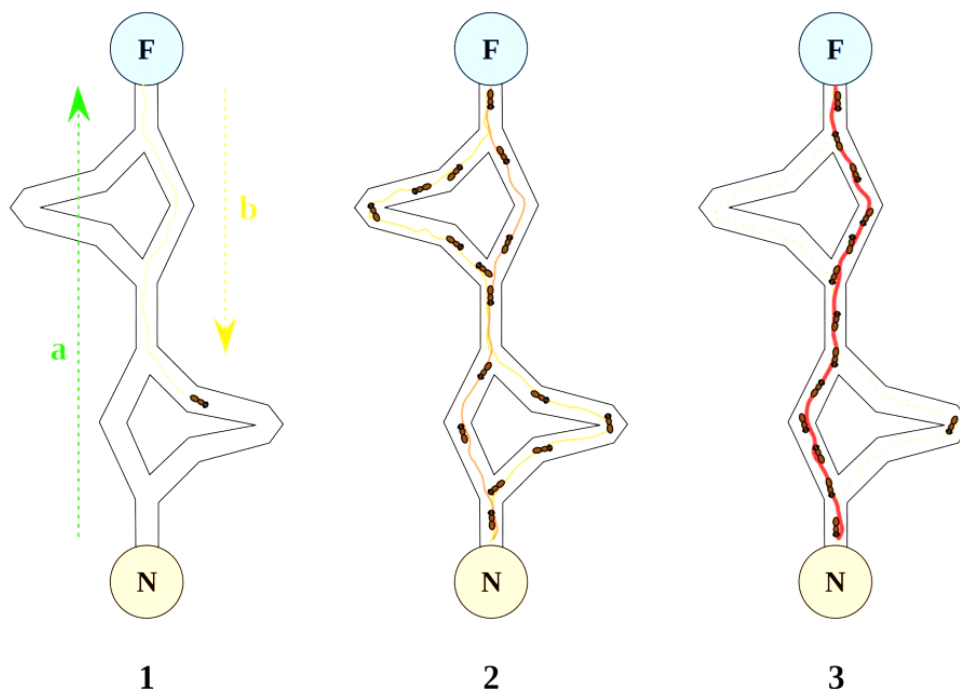
O Algoritmo Colônia de Formigas (ACO) é uma técnica de otimização que foi criado através das observações do comportamento coletivo das formigas quando saem a procura de fontes de alimentação (ALLEGRINI & OLIVIERI, 2011).

As formigas quando saem para encontrar comida se espalham de forma aleatória e retornam ao seu formigueiro (ninho). Quando elas encontram o menor caminho deixam um rastro de uma substância ao retornar a sua colônia chamado de feromônio (ALLEGRINI & OLIVIERI, 2011; PESSOA, 2015).

Essa substância biologicamente ativa faz com que quando as formigas encontram esse rastro. Elas tendem a não saírem mais de forma aleatórias em outros caminhos e tendem a seguir o caminho com feromônio. As formigas que encontram o formigueiro

primeiro, tendem a retornarem mais rapidamente ao ninho. Assim, a concentração dessa substância será maior naquele caminho percorrido com a menor distância alcançada (**Figura 6**). O feromônio evapora com o passar do tempo. Quanto maior a quantidade de formigas seguindo o rastro, maior a concentração do rastro dele deixado no caminho do trajeto, menor a probabilidade das formigas se desviarem no trajeto (ALLEGRIANI & OLIVIERI, 2011; PESSOA, 2015).

Esse algoritmo foi desenvolvido por Dorigo e Gambardella (1997). A sua primeira versão do ACO buscava solução de um problema de um caixeiro viajante, um problema de pesquisa de otimização combinatória no espaço de permutações. O fundamento do algoritmo é através da distribuição de  $m$  formigas em  $n$  cidades de forma a permitir que passe por cada uma das cidades e percorrendo o menor percurso para isso. Esse algoritmo é caracterizado através de duas formas: a primeira forma é de como é realizado a seleção das cidades e a segunda forma é a simulação da vaporização e de tal que reforço da marcação do feromônio.



**Figura 6:** Diagrama mostrando a desenvolvimento do tempo que as formigas levam no processo de busca do alimento e o seu o retorno para o ninho do levando o alimento. O ninho e a fonte de comida são otimizados por meio do trabalho conjunto da colônia (Adaptado: GOSS *et al.*, 1989).

A variação da intensidade de feromônio é realizada de forma que sua concentração em todas as trilhas ligando as cidades passe por dois processos de modificação:

- 1) Sua concentração é diminuída, devido à multiplicação por um fator entre 0 e 1, depois de cada iteração, desse modo, simula a evaporação do feromônio;
- 2) O aumento da concentração, toda vez que uma formiga utilizar a trilha.

A partir dessas duas condições, diversos algoritmos de otimização de colônia de formigas foram desenvolvidos para solucionar problemas de otimização, com a ideia de imitar o comportamento das formigas reais através das formigas “artificiais” (DORIGO, CARO & GAMBARELLA, 1999; SOCHA & DORIGO, 2008; MULLEN *et al.*, 2009; ALLEGRIANI & OLIVIERI, 2011; HEMMATEENEJAD *et al.*, 2011; RANZAN, 2014; PESSOA, 2015).

Podemos comparar a semelhanças das formigas reais com as artificiais:

**Caminho mais curto:** as formigas reais irão optar pelo caminho mais curto entre o alimento e o ninho, enquanto as artificiais, procuram o menor caminho a depender do problema a ser otimizado;

- **Colônia de agentes cooperativos:** as formigas agem de modo cooperativo no mundo real quanto nos virtuais. As reais fazem a deposição e evaporação do feromônio. As formigas artificiais são empregadas como uma variável matemática que conectará o processo;
- **Trilhas do feromônio:** nas formas de formigas o feromônio depositado pelas formigas atua da mesma maneira, modificando o ambiente, portanto, ratificando o aprendizado gerado pelas formigas;
- **Inteligência coletiva:** a inteligência é adquirida de modo coletivo nas duas formas de formigas, dado que, o comportamento individual é escasso ou aleatório;
- **Comportamento estocástico:** a forma probabilística é atributo nas duas realidades.

Porém, alguns atributos pertencem somente as formigas artificiais (DORIGO, CARO & GAMBARELLA, 1999):

- **Natureza do movimento:** o deslocamento das formigas artificiais é de maneira discreta, enquanto as formigas reais acontecem de modo contínuo;
- **Feromônio:** no algoritmo o depósito da substância feromônio, ocorre baseado na qualidade da solução encontrada;
- **Memória:** as formigas reais não possuem estrutura de memória como no caso das artificiais, que impossibilite de realizar movimentos.

O ACO é um algoritmo estocástico de investigação que possui dois tipos de convergências: valor e a solução. Quando se trata de problemas que há mais de uma solução ótima, o objetivo do algoritmo será convergir para qualquer um deles. Por isso, a convergência em valor avalia a probabilidade de que o algoritmo gere uma solução ótima pelo menos uma vez. Contudo, a convergência em uma solução analisa a probabilidade de que o algoritmo atinja o estado que continua gerando a mesma solução. Na otimização só é encontrado a solução ideal uma única vez, depois que encontra o problema é resolvido e o algoritmo pode ser interrompido (DORIGO & STÜTZLE, 2004).

Como relatado, a deposição do feromônio é uma alternativa de caminho para as formigas seguir. O algoritmo vê essa deposição como um valor inicial, que é um dos parâmetros para a primeira iteração que é constante, independentemente do número de formigas aplicado. Após a primeira iteração quando todas as formigas já percorreram o caminho, a quantidade de feromônio depositada será proporcional ao valor da função objetivo, sendo essa conta podendo ser realizado de diferentes modos. Contudo, a taxa deposição permanece constante, ou seja, a mesma quantidade de feromônio será depositada no caminho a cada iteração que foi percorrido para obter o melhor valor da função objetivo dentre as formigas da iteração (SOUZA, 2019).

Por meio de dois fatores, é realizado o cálculo de probabilidade de cada formiga (MULLHEN *et al.*, 2009):

- Fator heurístico: é calculado antes da primeira iteração e permanece inalterada no andamento do algoritmo;
- Fator associado ao feromônio: é atualizado a cada iteração.

O ACO pode ser descrito pela seguinte Equação:



$$P_m(c_k^i) = \frac{[Q \cdot \tau_k^i]^\alpha \cdot [\eta(c_k^i)]^\beta}{\sum_j [\tau_j^i]^\alpha \cdot [\eta(c_j^i)]^\beta} \quad (5)$$

onde  $c_k^i$  representa um componente da solução,  $\tau$  uma função do conjunto de solução que associa a cada um do componente a função objetivo,  $\eta$  uma função que está relacionado a cada componente da solução heurística,  $\alpha$  um valor positivo relacionado a quantidade de feromônio e  $\beta$  um valor positivo que se refere a informação heurística. Para os demais índices da equação refere-se  $m$  a formiga,  $i$  de iteração,  $k$  de algum possível elemento do conjunto solução e  $j$  referente a todos os elementos do conjunto solução. Usando dos valores da matriz de probabilidades, é aplicado o *random proportional rule*, ou seja, é o processo que aleatoriamente determina o conjunto de solução ponderado pelo valor das probabilidades calculadas. A probabilidade  $P_m(c_k^i)$  assume o valor 0, se essa transição para o modo  $ij$  infringir alguma restrição (DORIGO & BLUM, 2005). O ACO é aplicado para diversa áreas hoje em dia, como por exemplo, na robótica (LINGARA *et al.*, 2013), engenharia (SILVA, 2016), matemática (PIRES, 2019) e na quimiometria (RANZAN *et al.*, 2014; NACHE *et al.*, 2015).

### 3.2.5.2 Algoritmo Genético (GA)

No ano 1975, o pesquisador John H. Holland propôs a construção de um algoritmo matemático com base na teoria evolutiva das espécies para otimização de sistemas complexos (HOLLAND, 1975).

Esse algoritmo, chamado de Algoritmo Genético (GA), é um método que pretendia simular matematicamente todo o mecanismo da evolução biológica com todas as características e vantagens do processo. Com passar dos anos, o algoritmo passou a ser empregado na química para a resolução de diversos tipos de problemas (FILHO & POPPI, 1999).

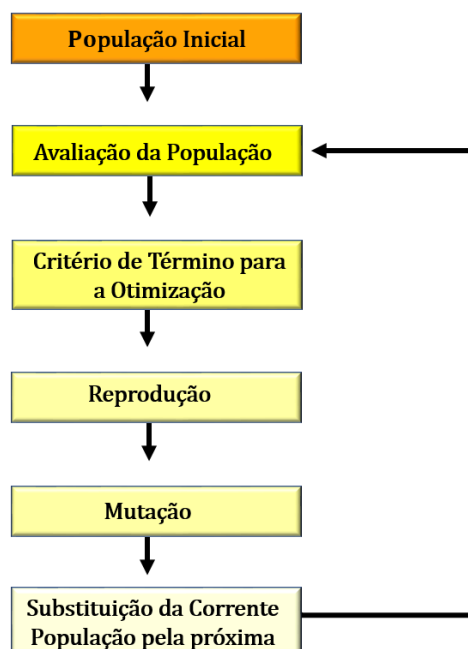
O algoritmo possui diversas vantagens quando aplicado no processo de otimização, como por exemplo (GOLDBERG, 1989):

- Não necessita de informação sobre gradiente da superfície de reposta;

- O desempenho da otimização não é afetado por eventuais descontinuidades das superfícies de reposta;
- A presença de mínimos locais não diminui o efeito do algoritmo;
- O desempenho do algoritmo tem apresentado ótimos resultados para os problemas de otimização de grande escala.

O GA assemelha-se muito ao processo evolutivo natural, pois ele foi estruturado de um modo que as informações referentes a um determinado sistema pudessem ser codificadas de maneira análoga aos cromossomos biológicos. O cromossomo biológico é composto por genes, que são responsáveis pelas características dos indivíduos. Por meio de uma analogia, foi possível construir um cromossomo artificial (FILHO & POPPI, 1999).

A base do GA envolve 5 etapas: codificação das variáveis, criação da população inicial, avaliação de reposta, cruzamento e mutação (**Figura 7**) (FILHO & POPPI, 1999).



**Figura 7:** Fluxograma representando as principais etapas do algoritmo genético básico (Adaptado: FILHO & POPPI, 1999).

Na primeira etapa o GA atua sobre uma população finita de elementos. Cada elemento é representado por um indivíduo codificado na configuração de estrutura cromossômica, ou seja, representa um potencial solução para algum problema específico. Então, cada gene apresentará um determinado parâmetro que será otimizado (MELANIE, 1998; FILHO & POPPI, 1999).

Logo, a criação de uma população inicial de cromossomos acontece por meio um gerador randômico, sendo que esta população segue de forma aleatória garantindo que não há nenhum tipo de influência tendenciosa externa (FILHO & POPPI, 1999).

No próximo passo que é a avaliação de reposta (aptidão), em cada etapa evolutiva do algoritmo, os indivíduos presentes na população são submetidos a avaliação por algum critério que tem como finalidade avaliar um valor de aptidão para cada indivíduo em relação ao objetivo a ser solucionado. Essa etapa da aptidão, é o resultado do procedimento mais importante do GA (MELANIE, 1998; FILHO & POPPI, 1999; NERY, 2017).

Por meio desse valor, os indivíduos que são classificados como aptos na população são selecionados através de um mecanismo que simula o processo de seleção natural (operador de seleção), para serem submetidos a operadores de transformação estocásticas inspirados em genética elementar e reprodução sexuada, com o objetivo de criar novos cromossomos mais aptos para a próxima geração (MELANIE, 1998; GEN, 2000).

Os operadores são classificados em 2 tipos: os operadores de cruzamento, que são responsáveis por gerar novos indivíduos através da combinação genética entre dois ou mais cromossomos progenitores e o outro o operador de mutação, que cria um novo indivíduo através de troca aleatória de algumas pequenas características genéticas de um indivíduo (MICHALEWICZ, 1996).

Para um desempenho satisfatório do GA, devemos observar a escolha apropriada da codificação dos cromossomos em conjunto com a definição de uma medida de aptidão, que deve avaliar criteriosamente os indivíduos da população. Por isso, a codificação correta dos indivíduos pode ajudar no maior número de soluções para o problema. A avaliação de modo eficiente dos indivíduos é importante para caracterizar quais cromossomos oferecem as soluções mais satisfatórias (HAUPT & HAUPT, 2004).

É muito importante que a medida de aptidão seja adequada, pois, define de maneira precisa quais cromossomos deverão possuir maiores probabilidades de serem selecionados para se reproduzirem, assim, irão transmitir suas características hereditárias as próximas gerações (NERY, 2017).

No GA, a cada geração, uma nova população de indivíduos, mais bem adaptadas ao problema, será reinserido no processo evolutivo. A interação entre avaliação, seleção de indivíduos, transformação estocástica e composição de uma nova população é repetido até que um critério de parada seja atingido. Esse critério pode variar do esgotamento do número finito de gerações pré-estabelecidas ou estagnação do processo de investigação em encontrar novas soluções com maiores aptidões (SILVA, 2011).

No processo de investigação, quando acontece a estagnação nesse processo, podemos dizer que o GA convergiu, e assim, como resultado, o ciclo evolutivo é encerrado e o algoritmo retorna o melhor indivíduo da população atual com a expectativa que ele seja próximo do valor satisfatório para o problema (MICHALEWICZ, 1996).

Quando são bem projetados, o GA tende a proporcionar resultados melhores para os problemas com espaço de busca muito extenso. Eles atuam sobre um conjunto de pontos, explorando múltiplas direções do espaço de busca simultaneamente e usam regras de transições probabilísticas nos operadores de transformação e seleção. (GOLDBERG, 1989; YANG, 2014).

Devemos ter um bom senso a respeito do algoritmo sobre a adaptação dos parâmetros de controle, particularmente relacionado ao tamanho da população de indivíduos, taxa de cruzamento e de mutação e critério de reinserção dos novos indivíduos na população de cromossomos (YANG, 2014).

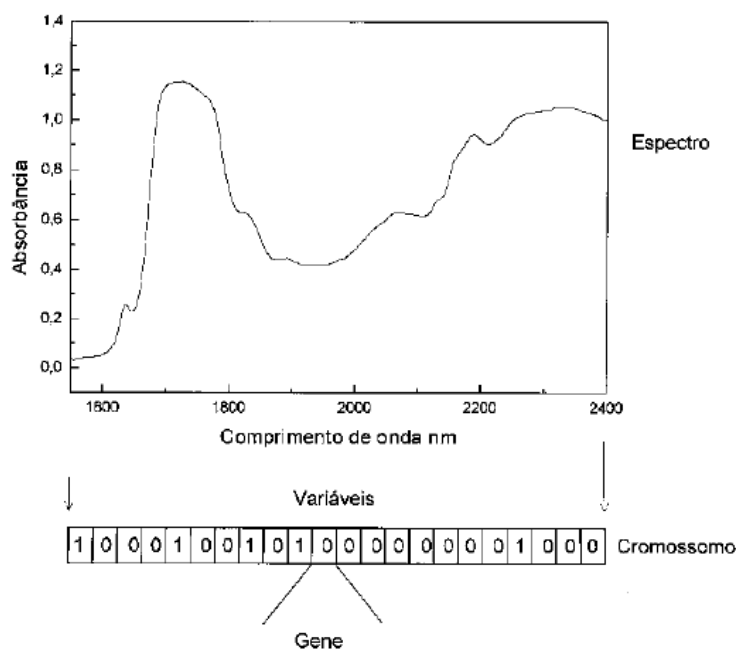
Esses parâmetros devem ser bem definidos pois, são extremamente importantes para o funcionamento do algoritmo, uma vez equivocado pode provocar um efeito indesejável no processo de busca e assim pode gerar resultados insatisfatórios. Não podemos esquecer, que assim como qualquer meta-heurística, os GAs não garantem encontrar excelentes globais para todos os problemas apresentados ou até mesmo em todos os casos de um determinado problema. Contudo, a aplicação desse algoritmo tem se mostrado um papel eficiente na busca de soluções competitivas em problemas complexos (GOLDBERG, 1989; MELANIE, 1998; YANG, 2014).

Na química o GA é muito empregado em diversos métodos para aplicação de inúmeros problemas complexos de otimização de sistemas essencialmente devido a sua robustez e versatilidade. Todavia, a aplicação mais difundida do GA está na relacionado a seleção de variáveis principalmente na área da espectroscopia (BANGALORE *et al.*, 1996; FILHO & POPPI, 1999).

Em algumas técnicas analíticas muitas aplicações da calibração multivariada envolvem um enorme número de variáveis. Na grande maioria das vezes, o uso somente

de algumas variáveis que contêm mais informações podem ocasionar maior segurança e praticidade na interpretação do modelo desenvolvido. De tal modo, as variáveis que têm uma menor relação sinal/ruído ou/e não linearidade pode ser descartadas (FILHO & POPPI, 1999).

No processo de aplicação do GA para a seleção de variáveis espectrais, temos a codificação do problema, considerando-se que o cromossomo possui “p” genes, onde cada gene representa uma variável do espectro (comprimento de onda). Portanto, o cromossomo terá o mesmo número de variáveis contidas no espectro (**Figura 8**) (FILHO & POPPI, 1999).



**Figura 8:** Representação de uma codificação espectral na forma de um cromossomo. Neste caso, cada valor de comprimento corresponde a um gene do cromossomo (Adaptado: FILHO & POPPI, 1999).

Assim, ao utilizar o GA para a seleção de variáveis, é necessário, dependendo do modelo de calibração, empregar determinadas “condições de contorno”, como por exemplo, pré-estabelecer o número máximo de variáveis selecionadas, como no caso da regressão linear múltipla (FILHO & POPPI, 1999).

No GA, o código binário (0,1) é empregado na seleção de variáveis para o auxílio na codificação de problemas. Portanto, cada gene pode admitir o valor *um* ou *zero*, sendo que quando a posição referente a uma determinada variável for igual a *um*, isto implica

na seleção desta variável. Por outro lado, se a posição contiver valor **zero**, isto indica que a variável não foi selecionada (FILHO & POPPI, 1999).

### 3.2.5.3 Algoritmo *Stepwise* (SW)

O método *stepwise* seleciona subconjuntos de variáveis através da busca exaustiva a partir de uma função custo que visa melhorar o desempenho dos modelos, seja ele de regressão ou de classificação. Na estratégia utilizada o mesmo pode adicionar (*forward selection*) ou remover (*backward elimination*) uma variável por vez, ou ainda combinar os dois procedimentos, buscando a construção do melhor modelo (BROWN, TAULER, R & WALCZAK, 2009).

Em 2005, Galvão e colaboradores apresentaram uma versão para a seleção de amostras na construção de modelos de regressão por mínimos quadrados parciais, comparando este método com o clássico Kennard Stone, bem como com a seleção randômica (GALVÃO *et al.*, 2005). No ano seguinte, Caneca e colaboradores propõem a aplicação deste método para a seleção de variáveis de dados espectrais obtidos no infravermelho médio e próximo, buscando a construção de modelos de classificação e regressão multivariada (CANECA *et al.*, 2006).

O algoritmo proposto avalia o valor individual de cada variável espectral de acordo com sua discriminabilidade em relação às classes em consideração (DUDA, HART & STORK, 2001).

Em problemas de classificação, as variáveis podem ser classificadas com base em sua capacidade de discriminar as classes em questão. De acordo com Duda *et al.*, a discriminabilidade  $D_i$  da variável  $x_i$  pode ser quantificada conforme **Equação 6**.

$$D_i = \frac{S_{Bi}}{S_{wi}} \quad (6)$$

onde  $S_{wi}$  e  $S_{Bi}$  são medidas das dispersões dentro e entre as classes para a variável  $x_i$ , respectivamente. A dispersão dentro da classe  $S_{wi}$  é definida conforme **Equação 7**.

$$S_{wi} = \sum_{j=1}^C s_{ij} \quad (7)$$

onde  $s_{ij}$  é a dispersão de  $x_i$  na classe  $j$ , calculada conforme **Equação 8**.

$$s_{ij} = \sum_{k \in I_j} [x_i^k - m_{ij}]^2 \quad (8)$$

onde  $x_i^k$  denota o valor de  $x_i$  no  $k$ -ésimo objeto e  $m_{ij}$  é o valor médio de  $x_i$  na classe  $j$ , ou seja:

$$m_{ij} = \frac{1}{n_j} \sum_{k \in I_j} x_i^k \quad (9)$$

Já a dispersão entre classes  $S_{Bi}$  é definida conforme **Equação 10**.

$$S_{Bi} = \sum_{j=1}^c n_j [m_{ij} - m_i]^2 \quad (10)$$

onde  $m_i$  é a média de  $x_i$  sobre todas as amostras de treinamento.

Em cada etapa, a variável  $x_i$  com a maior discriminabilidade  $D_i$  é selecionada, um procedimento de validação cruzada *leave-one-out* é realizado e o número de erros é computado. Antes da próxima etapa, as variáveis altamente correlacionadas com as já selecionadas são descartadas para evitar problemas de colinearidade. O algoritmo para quando não há mais variáveis disponíveis. O conjunto de variáveis que resultou no menor número de erros de validação cruzada é então utilizado para a construção do modelo de classificação desejado.

Os principais passos desta estratégia são descritos a seguir:

Sejam  $A$  e  $B$  os conjuntos de índices para as variáveis já selecionadas e aquelas ainda disponíveis, respectivamente. Além disso, seja  $L$  um limite de correlação ( $0 < L < 1$ ) parâmetro de entrada a ser definido pelo analista. A seguir,  $N$  é um contador que indica o número de variáveis já selecionadas.

- **Etapa 0 (inicialização).**  $A = \{\}$ ,  $B = \{1, \dots, d\}$ ,  $N = 0$ .
- **Etapa 1.** Calcule  $D_i$  para  $1 \leq i \leq d$ .
- **Etapa 2.**  $i^* = \arg \max D_i, i \in B$ .
- **Etapa 3.** Mova  $i^*$  de  $B$  para  $A$ . Seja  $N = N + 1$ .
- **Etapa 4.** Realize um procedimento de validação cruzada *leave-one-out* usando as variáveis com índices em  $A$ . Seja ECV ( $N$ ) o número de erros de validação cruzada resultantes.

- **Etapa 5.** Calcule o coeficiente de correlação múltipla  $r_i$  de cada variável  $x_i$  com índice em  $B$  em relação às variáveis com índices em  $A$ .
- **Etapa 6.** Exclua de  $B$  os índices de variáveis com coeficiente de correlação múltipla maior que  $L$ .
- **Etapa 7.** Se  $B = \{\}$ , volte para a Etapa 2.
- **Etapa 8.** O número ótimo  $n^*$  de variáveis é obtido a partir do mínimo de ECV ( $n$ ),  $n = 1, \dots, N$ . As variáveis selecionadas correspondem aos primeiros  $n^*$  índices em  $A$ .

### 3.3 Matriz Vinho Tinto

As propriedades das uvas vermelhas apresentam diversos efeitos na dieta humana. Nesta seção, será apresentada uma breve revisão bibliográfica de alguns aspectos da história das descobertas das uvas e dos vinhos no mundo e no Brasil, até a chegada à região Sul. Alguns dados sobre a vitivinicultura no mundo e no Brasil também são apresentados a seguir, além das diferenças das duas regiões estudadas: Serra Gaúcha e Campanha Gaúcha. Por fim, é abordada a composição e o pH do vinho tinto e seus benefícios à saúde.

#### 3.3.1 Vinho Tinto

#### 3.3.2 História de Uvas e Vinhos

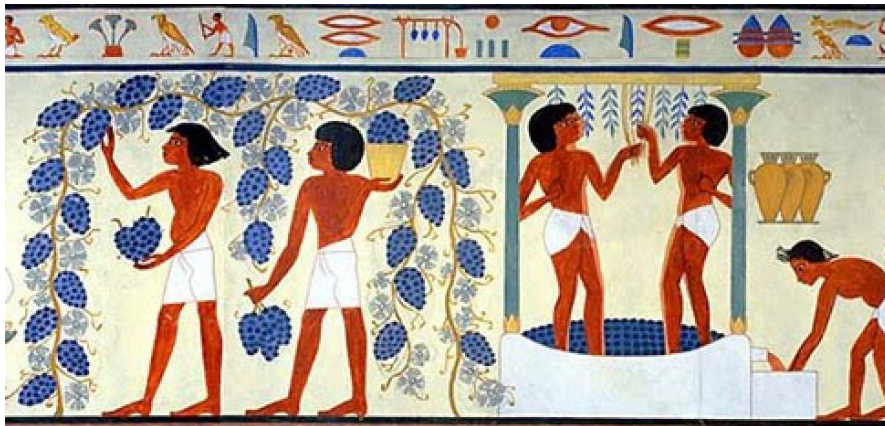
As videiras são arbustos herbáceos que pertencem à família *Vitaceae*, cujo fruto é a uva (*Vitis spp*) (GUERRA *et al.*, 2005). Não se sabe apontar precisamente o local e a época em que o vinho surgiu. Em condições edafoclimáticas, a europeia é a mais sensível e algumas variedades requerem um manejo cuidadoso dos frutos para uma produção com alta qualidade (KUHNS *et al.*, 1996).

A espécie de videiras americanas (ex. *Vitis labrusca* L.) tolera um manejo menos rigoroso e tem maior resistência às doenças quando comparada com a europeia. Sua qualidade sensorial é apropriada para o consumo *in natura* e também para a elaboração de sucos e vinhos de mesa (KUHNS *et al.*, 1996).



Os primeiros indicadores de elaboração de vinhos são datados em 4.000 a.C., no Egito, porém acredita-se que o vinho possa ter surgido antes dessa época e somente nesse período ocorreram os registros em pinturas (GASNIER, 2015).

Nas tumbas dos faraós foram encontradas pinturas (**Figura 9**) com os detalhes de diversas etapas do processo de elaboração do vinho, tais como: a colheita da uva, prensagem e a fermentação. Várias cenas nessas pinturas mostravam o consumo da bebida em taças, jarras e canudos, nos ambientes festivos elegantes. O vinho parece ter sido ofertado apenas para ricos, nobres e sacerdotes. Alguns registros mostram o vinhedo e o vinho sendo ofertados aos deuses, principalmente pelos faraós (JOHNSON, 2004).



**Figura 9:** Representação da elaboração de vinho registrada pelos egípcios através de pinturas no ano de 1100 a.C. (Adaptado: PRATES, 2018).

Escavações em vários países revelaram a existência de sementes de uvas da Idade da Pedra, cerca de 8.000 a.C. Os arqueólogos aceitam o acúmulo de sementes de uvas como evidência de elaboração de vinhos. A popularização do consumo e o desenvolvimento de práticas enológicas iniciaram a partir do desenvolvimento das civilizações grega e romana (GASNIER, 2015).

Alguns pesquisadores encontraram na região do Irã vestígios de vinhos em ânforas que datam de 5000 anos atrás, muitos acreditam que nessa região iniciou a produção do vinho. Porém, a teoria mais aceita nos dias de hoje é que a uva teria se originado no Cáucaso, chegando depois ao Oriente Médio (LILLA, 2016).

Sabemos que o vinho existe muito antes mesmo da escrita, desde os primeiros textos sumérios. Esses textos foram gravados em argila e contam uma história semelhante ao

dilúvio, narrado nas Sagradas Escrituras. Segundo o relato, Gilgamesh, que era um rei, sumério recebeu ordens para construir uma arca no meio do deserto pagando os trabalhadores com cerveja e vinho. A Bíblia narra que, após o dilúvio, Noé saiu da arca para plantar uvas e produzir vinhos. Existem outros relatos também narrados nas Escrituras, como por exemplo, quando Jesus Cristo operou o seu primeiro milagre, após um pedido de sua mãe nas Bodas de Caná da Galileia, transformando então, água em vinho e, também, a utilização do vinho na Santa Ceia, em que Jesus afirma ser seu sangue (LILLA, 2016).

Na mitologia grega e nas lendas persas, o vinho também foi muito importante. Era consumido no dia a dia pelos gregos e romanos; os gregos difundiram as videiras pela região mediterrânea até a Espanha. Na “Magna Grécia”, sul da Itália, a produção progrediu tanto, que na época a região foi chamada de Enótria, “terra do vinho”. Depois de incentivarem vários povos, o império romano começou um estudo sério de assuntos relacionados ao vinho: as uvas, modo de plantar e métodos de produção (LILLA, 2016).

Na idade Média, a Europa passou por grandes dificuldades, afetando também, a produção de vinho. Os mosteiros eram a única exceção; eles se tornaram centros de cultivos e produção de vinho. Os monges faziam toda a pesquisa relacionada à bebida e, com isso, a igreja se tornou proprietária dos maiores e mais importantes vinhedos da Europa (LILLA, 2016).

No período da Renascença a qualidade do vinho evoluiu bastante. Devido às grandes viagens marítimas realizadas na época, mudas de uvas eram levadas para o plantio em novas terras, pois as uvas que havia na Europa não eram adequadas à produção de vinhos de qualidade. Somente com a Revolução Industrial, começou a fabricação de uma série em garrafas de vidro e rolhas de cortiça e, a partir do século XVIII e XIX, o vinho foi se transformando no que conhecemos hoje (LILLA, 2016).

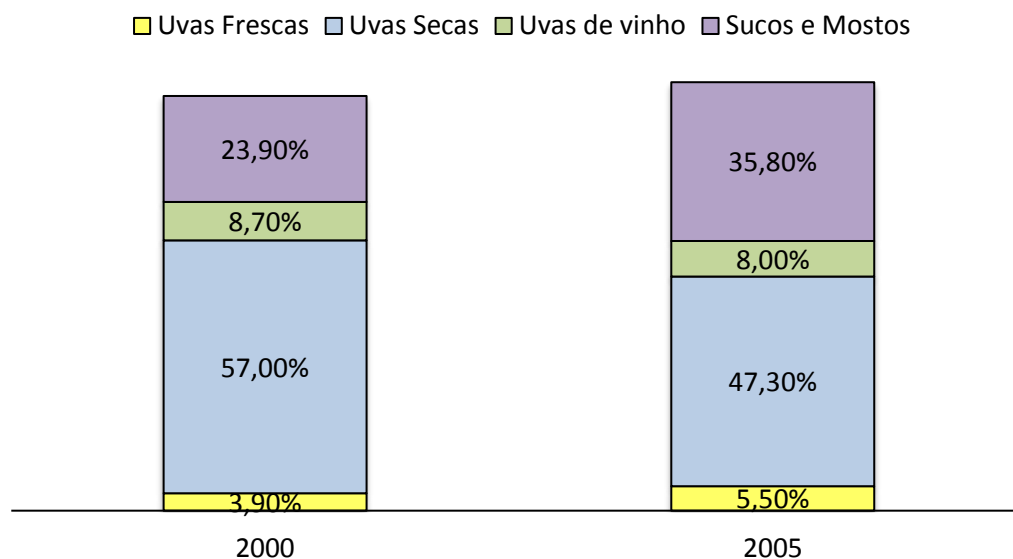
Nessa mesma época, cientistas traziam novidades para o processo de produção de vinho, principalmente Lavoisier, que estudou a fermentação, e Chaptal, ministro de Napoleão, que criou o processo de adição de açúcar ao mosto, para elevar o grau alcoólico, melhorando assim a conservação (que é chamado de *chaptalização*). O cientista mais importante foi Pasteur, que foi o responsável por explicar cientificamente o processo que ocorre na fermentação, visto que, mesmo sendo realizado esse processo na prática, ninguém sabia como acontecia (LILLA, 2016).

Nessa fase, surgiu uma praga de inseto filoxera (espécie *Daktulosphaira vitifoliae*) na Europa, que acabou com os vinhedos no fim do século XIX e destruiu a produção de vinhos no continente. Os enólogos começaram a surgir após a Segunda Guerra Mundial e retomaram a produção do vinho de forma mais científica. Viagens e descobertas de novas terras, o vinho foi levado a diversas regiões. Sendo assim, houve uma adaptação em regiões mais quentes devido ao aperfeiçoamento o controle de temperatura em lugares com temperaturas mais altas (LILLA, 2016).

### 3.3.3 Vitivinicultura Dados Macroeconômicos

A uva e o vinho abrangem mais de 40 países nos cinco continentes, conforme a Organização Internacional de Vinha e Vinhos (OIV) (EMBRAPA, 2015). A OIV é uma organização intergovernamental de natureza científica e técnica, de competência reconhecida nos domínios da videira, vinho, bebidas à base de vinho, uvas de mesas, uva passa e outros produtos da videira. Anualmente publica um relatório estatístico sobre Vitivinicultura Mundial. Segundo o relatório referente ao ano de 2016, a área de vinhedo, em produção ou em fase de produção a nível global de uvas para vinho, uvas de mesa ou uvas secas, os cinco maiores produtores são a Espanha (14%), China (11%), França (10%), Itália (9%) e Turquia (7%) (OIV, 2017).

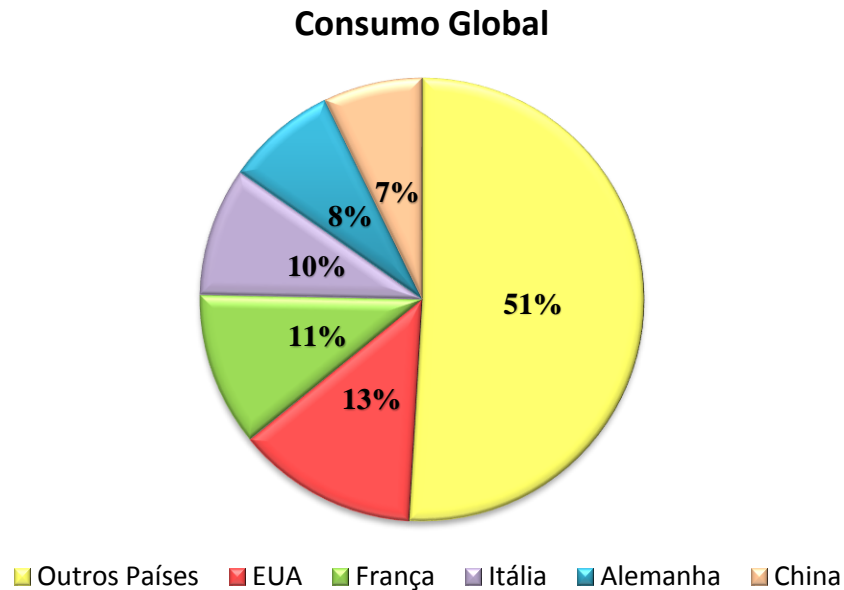
O Brasil aparece na 19<sup>o</sup> colocação nessa lista. Em nível global, a produção de uva na Europa é responsável por 39% da produção mundial, seguida pela Ásia com 34% e América com 18%. Podemos observar na **Figura 10**, a produção total mundial de acordo com os tipos de produtos, entre o ano de 2000 e 2005 (OIV, 2017).



**Figura 10:** Gráfico da produção total mundial de acordo com os tipos de produtos (Adaptado: OIV, 2017).

Na produção total de uvas relacionada ao ano de 2005, os países como a China, Itália e Estados Unidos aparecem no topo da lista. Já o Brasil, aparece como 16º, com produção de 67% das uvas frescas e 33% de uvas de vinho (OIV, 2017).

Na produção mundial de vinhos do ano de 2016, os maiores produtores foram a Itália, seguida da França, Espanha e Estados Unidos. O Brasil está na 20ª colocação do ranking. Referente aos países que mais exportam vinho no mundo, os países do Velho Mundo ainda são os que mais se destacam. Espanha, Itália e França, por exemplo, exportam mais da metade de todo o vinho consumido no mundo. Já no ranking do consumo global dos vinhos, os países que lideram: são os Estados Unidos em primeiro lugar, seguido da França, Itália, Alemanha e China, conforme mostra a **Figura 11**. O Brasil fica na 18ª posição do consumo global (OIV, 2017).



**Figura 11:** Gráficos dos cinco países que consomem metade do vinho do mundo (OIV, 2017).

Portugal é o primeiro país do mundo se o consumo anual de vinho per capita for levado em conta. A produção de vinhos no país é muito intensa devido ao seu clima temperado e úmido, com estação seca no verão e temperaturas amenas no inverno, porém cada região possui sua própria identidade (CLIMACO *et al.*, 2012). Cada português com idade acima de 15 anos bebe, em média, 54 litros por ano. Os países que seguem no ranking de consumo per capita são: a França (51,8 litros), a Itália (41,5), a Suíça (40,4) e a Áustria (32,4). Segundo o estudo da OIV, o consumo per capita de vinhos no Brasil é de cerca de dois litros e vem se mantendo estável na última década, ocupando assim a 20ª colocação (OIV, 2017).

#### 3.3.4 História da Vitivinicultura no Brasil

Existem várias espécies de videiras no mundo, classificadas em europeias, americanas e híbridas, porém, no Brasil, a espécie europeia (*Vitis vinífera* L.) é a mais utilizada na produção de vinhos finos (GUERRA *et al.*, 2005).

Os registros históricos apontam que as primeiras videiras no Brasil foram trazidas pelas expedições dos colonizadores portugueses, no ano de 1532, através de Martin

Afonso de Souza. As mudas foram plantadas na Capitania de São Vicente no sudeste do país, hoje o estado de São Paulo. As condições desfavoráveis de clima e solo da região impediram que a experiência avançasse (BOTELHO & PIRES, 2009).

Brás Cubas, fundador da cidade de Santos, era um membro da expedição colonizadora de Martin Afonso de Souza que insistiu no cultivo de videiras transferindo suas plantações para o litoral ao Planalto Atlântico. Assim sendo, é reconhecidamente, o primeiro a cultivar a vinha em nossas terras. Somente no ano de 1551, ele consegue elaborar o primeiro cultivo de vinha brasileira (ABE, 2018).

Em seguida, a Viticultura expandiu-se para as outras regiões do País, sempre com os cultivares *Vitis viníferas* procedentes de Portugal e da Espanha (BOTELHO & PIRES, 2009; IBRAVIN, 2018a).

No Sul do Brasil, a viticultura foi impulsionada pela chegada dos Jesuítas à região. No ano de 1626, o jesuíta Roque González de Santa Cruz começou o cultivo de videiras europeias no estado do Rio Grande do Sul, mais precisamente, em São Nicolau, nos Sete Povos das Missões. Roque era padre e contava com a ajuda dos índios para a elaboração dos vinhos e elementos das celebrações religiosas. Porém, as variedades de viníferas tinham muita dificuldade de adaptação nessas terras, fazendo com que impedisse a disseminação da viticultura no Brasil (REAL, 1981; DARDEAU, 2015; ABE, 2018).

Apesar do Sudeste, na época, não ser um local adequado para o cultivo de uvas, os produtores eram persistentes. No ano de 1640, foi realizada a primeira degustação orientada no Brasil, conforme relata a 1<sup>o</sup> Ata da Câmara de São Paulo. A intenção era padronizar os vinhos produzidos no país (PIEROZAN, 2017, IBRAVIN, 2018a).

No ano de 1732, a vitivinicultura rio-grandense começa a renascer com a chegada dos imigrantes portugueses compostos por sessenta casais açorianos e madeirenses radicados em Rio Grande, Pelotas e Porto Alegre. Apesar de trazerem mudas das ilhas dos Açores e da Madeira, as plantações não ganharam muita expressão (REAL, 1981; ABE, 2018; IBRAVIN, 2018a).

Em 1789, devido às diversas iniciativas de vinicultura no Brasil, a corte portuguesa proibiu o cultivo de uva no país como forma de proteger sua própria produção. Com isso, houve redução da comercialização da bebida na colônia e a atividade ficou restrita ao meio doméstico. Mas, em 1808, com a transferência da coroa Portuguesa para o Brasil, ocasionando a vinda da família real essa proibição foi anulada (REAL, 1981; PIEROZAN, 2017).

Oficialmente, o plantio de videiras e a produção de vinho no Rio Grande do Sul acontecem em 1813, por Manoel de Macedo Brum da Silveira, na cidade de Rio Pardo e é reconhecido por D. João VI. A partir desse período até 1835, ele registra a elaboração que rendeu a primeira carta-patente para a produção da bebida no país. O crescimento da produtividade intensificou ainda mais com a colonização alemã que possuía grandes interesses pelo vinho. Nesse mesmo período, o italiano João Batista Orsi estabelece na Serra Gaúcha e inicia o cultivo de uvas europeias (REAL, 1981).

Somente em 1840, por meio do inglês Thomas Messiter, foram introduzidas outras espécies de uvas no estado do Rio Grande do Sul, a saber, a *Vitis labrusca* e a *Vitis bourquina*. A variedade de uva americana conhecida como Isabel, da espécie *Vitis labrusca*, foi implantada na Ilha dos Marinheiros, na Lagoa dos Patos, e teve destaque com sua grande adaptação nas nossas terras por ser resistente às doenças. Depois disso, por volta de 1860, ela se espalhou por todo o estado do Rio Grande do Sul cativando os agricultores que a cultivavam (REAL, 1981).

Com os imigrantes italianos, que chegaram no Brasil ao ano de 1875, a vitivinicultura deu um grande salto, pois eles traziam da sua terra natal todo o conhecimento e técnicas de elaboração do vinho, além da cultura do consumo da bebida; sendo assim, agregaram uma importância econômica à atividade. Em 1881, houve o primeiro registro de elaboração de vinho no Vale dos Vinhedos, localizado na cidade de Garibaldi, Rio Grande do Sul, com uma produção total de 500 mil litros segundo os registros (REAL, 1981).

Uma situação desfavorável para a produção de uvas na Serra Gaúcha era o excesso de chuvas. Em 1908, com a chegada dos enólogos Lourenço e Horácio Mônaco, os conhecimentos dos imigrantes italianos foram agregados aos deles possibilitando soluções para a produção de videiras nas áreas chuvosas (DARDEU, 2015).

No ano de 1929, durante um período de aproximadamente 10 anos, 26 cooperativas foram fundadas, sendo que muitas existem até hoje. Isso resultou na fundação do sindicato do vinho com a intenção de organizar o setor mediante o crescimento da atividade. Na década de 70, empresas estrangeiras começam a ter interesse no Brasil, ampliando assim, o cultivo da uva e as novas técnicas (IBRAVIN, 2018a). Desde a colonização, a produção de vinhos no Brasil não parou, sendo o maior país da América Latina e o quinto maior produtor vitivinícola do hemisfério sul (REAL, 1981).

No Brasil, a área cultivada de videiras no ano de 2016 foi de 77.786 hectares, sendo a maior concentração no estado do Rio Grande do Sul com 64,30% da área vitícola nacional. A produção de vinhos, sucos e derivados de uva no Rio Grande do Sul foi de 244,92 milhões de litros neste mesmo ano, tendo o Brasil produzido neste período 984,44 toneladas de uvas (MELLO, 2017).

No Brasil, existem diversos estados que cultivam uvas para a produção de vinhos, a saber, o Rio Grande do Sul, São Paulo, Santa Catarina, Paraná, Minas Gerais, além da Bahia e Pernambuco, no Vale do São Francisco (IBRAVIN, 2018b). O estado de Pernambuco possui o segundo lugar do ranking entre os maiores produtores de uvas, seguido de Paraná, Santa Catarina, Bahia, Minas Gerais, Goiás, Paraíba e Espírito Santo. Os demais estados somam uma participação de 4,5% de volume (t) total produzido (IBGE, 2017).

A vitivinicultura no Brasil está espalhada em vários Estados, porém se concentra em poucas regiões. No Rio Grande do Sul, tem grande abrangência na Serra Gaúcha, onde quase toda produção se destina à agroindústria do suco e do vinho e essencialmente é realizada pelos agricultores de agricultura familiar. No Vale do São Francisco (Pernambuco e Bahia) e em São Paulo, a renda para milhares de famílias é proveniente da produção de uvas de mesa (EMBRAPA, 2015).

O Estado de Santa Catarina também produz vinhos finos nas regiões de São Joaquim, junto ao Planalto Sul Catarinense. Nessas regiões há vinhedos de altitude, sendo que São Joaquim representa a região produtora de vinho. No Nordeste, as colheitas de uvas acontecem ao longo do ano, já na região Sul, colhe-se uma safra por ano (EMBRAPA, 2015).

A implantação das indicações Geográficas no Brasil, vem crescendo nos últimos anos; a viticultura tem colaborado intensamente para o desenvolvimento dos territórios envolvidos, promovendo assim a agregação de valores aos produtos e a valorização de seus fatores naturais e culturais (EMBRAPA, 2015). Todas as regiões possuem diferentes características de clima, solo, variedades de uvas, sistemas de produção dos vinhos desde a vinificação ao envelhecimento, que possibilitam uma grande diversidade de aroma e sabor dependendo das características peculiares de cada região. Estes aspectos constituem a qualidade da vitivinicultura brasileira (EMBRAPA, 2005).



### 3.3.5 Vitivinicultura no Estado do Rio Grande do Sul

Ainda na safra de 2018, 663,2 milhões de quilos de uvas foram destinados ao processamento de produtos vinícolas no Rio Grande do Sul. Do total, 597.699.541 quilos foram de uvas americanas e híbridas e 65.540.421 quilos da espécie *Vitis viníferas* (IBRAVIN, 2018b).

Nesta safra, 113 variedades de uvas foram colhidas em 129 municípios do Rio Grande do Sul, com processamento realizado em 64 cidades do estado. Seguindo o padrão dos últimos anos, 50% da produção foi destinados para a elaboração do suco de uva. As principais regiões produtoras de sucos e vinhos são: Serra Gaúcha, Campos de Cima da Serra, Campanha, Serra do Sudeste (IBRAVIN, 2018b).

#### 3.3.5.1 Vitivinicultura na Serra Gaúcha

A maior parte da produção de uvas e vinificações no Brasil ocorre no estado do Rio Grande do Sul, mais especificamente na Serra Gaúcha (EMBRAPA, 2015, IBRAVIN, 2018b; MELLO, 2013).

A Serra Gaúcha é considerada a principal região vinícola do Brasil. A agricultura familiar tem uma grande participação e é representada por cerca de 12.000 pequenas propriedades que cultivam 31.000 hectares de vinhas. Possui aproximadamente 600 produtores de vinhos entre cooperativas, grandes empresas e vinícolas familiares, que ao todo somam 350 milhões de litros industrializados por ano (NIEDERLE, 2009).

Localizada no nordeste do Rio Grande do Sul, a Serra Gaúcha é uma região montanhosa, de clima úmido cortada pelo paralelo 29°. Possui uma altitude que pode variar entre 600 e 1000 m, com índice de precipitação de cerca de 1.800 mm/ano, com solos ácidos e arenosos, ricos em matéria orgânica e com drenagem pouco eficiente. Durante os meses de colheita, a Serra Gaúcha sofre com os problemas causados pelo excesso de chuva. As uvas produzidas nessa região são ideais para o preparo do vinho-base utilizado na elaboração de espumantes, pois os métodos de plantio e o clima favorecem a produção regularmente de uvas com alta acidez (FARIAS, 2014).

A Serra Gaúcha pode ser dividida por três microrregiões: Vale dos Vinhedos, Pinto Bandeira e Altos dos Montes. As principais castas tintas que são produzidas nesses locais são: Merlot, Cabernet Sauvignon, Tannat (FARIAS, 2014).

Em 2001, com área de aproximadamente 8.000 ha, o Vale dos Vinhedos foi a primeira área demarcada do Brasil com indicação de procedência “Vale dos Vinhedos”; os municípios que demarcam essa região são Garibaldi, Bento Gonçalves e Monte Belo do Sul (FARIAS, 2014).

### 3.3.5.2 Vitivinicultura na Região da Campanha

A Campanha é uma região localizada no sudeste do Rio Grande do Sul. Sendo local de fronteira entre Brasil e Uruguai. A linha de divisão entre esses países é marcada com campos nativos formando coxilhas (CHELOTTI, 2006).

A região é muito jovem em termos de vitivinicultura quando comparada às demais regiões que possuem mais de 100 anos de tradição, o que explica seu baixo reconhecimento nesse ramo de atuação (CERSÓSIMO & TECHEMAYER, 2020). Porém, nas últimas décadas a Campanha vem expandindo a vitivinicultura para produção de vinhos e espumantes (SARMENTO, 2016).

A Campanha é muito ampla, pois envolve cerca de 20 municípios segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), porém os municípios que possuem atividade vitivinícola são apenas 10, a saber: Alegrete, Quaraí, Itaqui, Rosário do Sul, Santana do Livramento, Dom Pedrito, Hulha Negra, Bagé, Candiota e Uruguaiana (FLORES, 2011; FARIAS, 2014).

Atualmente existem 18 empresas na região da Campanha com atividade vitivinícola (FLORES, 2011; DARDEAU, 2015). A produção vitivinícola dessa região teve um grande impulso na década de 80 devido ao grande investimento e expansão das vinícolas da Serra Gaúcha para a Campanha. Isso levou à fundação da Associação dos Produtos de Vinhos Finos da Campanha Gaúcha com o objetivo da busca por uma indicação geográfica, uma vez que cada território possui certa complexidade do *terroir* (FLORES, 2011; FLORES & MEDEIROS, 2013).

A Campanha situa-se no paralelo 31° e apresenta topografia plana com índice pluviométrico de 1.400 mm/ano. É uma região parcialmente de clima seco e, dependendo do ano, há necessidade de irrigação no período do verão que pode chegar a 35°C. Possui boa amplitude térmica podendo alcançar 13°C. A Campanha possui solo com baixa acidez, arenoso e argiloso com boa drenagem e baixa matéria orgânica, e, normalmente tem clima estável entre as safras (FLORES, 2011; FLORES & MEDEIROS, 2013);

FARIAS, 2014). A característica plana favorece excelentes condições de luminosidade que é desejável para a produção de vinhos finos com qualidade (GUERRA *et al.*, 2009).

A variação de temperatura na Campanha, com dias quentes e noites frias, faz com que a maturação das uvas seja lenta e, conseqüentemente, possuam maiores teores de açúcares e maiores quantidades de polifenóis (SOUZA *et al.*, 2006).

Outra característica observada nos vinhos finos elaborados com uvas dessa região é a intensidade do corpo e da cor, que está relacionado com a quantidade de composto fenólicos presente na composição química (GUERRA *et al.*, 2009). As principais castas tintas produzidas nessa região são: Merlot, Cabernet Sauvignon, Catas Portuguesas, Tannat, Pinot Noir (FARIAS, 2014).

### 3.3.6 Composição e pH do Vinho

Os compostos fenólicos estão presentes nas frutas e vegetais, sendo as uvas uma das maiores fontes (MAXCHEIX *et al.*, 1990). Existe uma grande diversidade entre os cultivares de uvas, possibilitando diferentes características na fruta, tanto de sabor quanto de coloração (ABE *et al.*, 2007).

Os compostos fenólicos, ou simplesmente fenóis, são responsáveis pelas principais diferenças entre os vinhos brancos e tintos, relacionadas com a cor e também com o sabor dos produtos (RIBERÉAU-GAYON *et al.*, 2006). Os compostos fenólicos encontram-se em maior quantidade nos vinhos tintos, com concentrações presentes nas cascas e sementes das uvas (GUERRA, 2010).

As uvas são a matéria-prima para elaboração e produção de vinhos e sucos, sendo assim, é fundamental conhecer os compostos fenólicos, uma vez que estes podem influenciar na qualidade dos produtos finais. A coloração do vinho tinto é outro fator que pode influenciar na qualidade, pois é a primeira característica com a qual os consumidores têm contato e está diretamente associada à qualidade do produto (ABE *et al.*, 2007).

Além disso, fatores como, variedade da uva, maturidade na colheita, tipo de solo, condições climáticas e processos de vinificação podem influenciar na composição do vinho fazendo com que cada bebida tenha suas características próprias (SAAD, 2016).

O vinho é uma bebida resultante da fermentação alcoólica completa ou parcial da uva fresca, esmagada ou não, ou do mosto simples ou virgem, com um conteúdo alcoólico mínimo de 7% (V/V a 20° C) (BRASIL, 1988); possui um pH que normalmente varia

entre 3,2 e 4,0 e é uma matriz complexa que pode sofrer bastante variações, contendo açúcares, ácidos, compostos voláteis, fenólicos, e uma variedade de produtos químicos como, compostos peptídicos, proteínas, carboidratos, tióis, além de ser uma solução que varia entre 10% a 15% de etanol em água (JACKSON, 2008a).

Os vinhos são misturas de ácidos fracos e quando combinados formam sais em maior ou menor grau de acordo com sua constante de acidez ( $K_a$ ). Essa proporção de sais varia de acordo com a origem geográfica, variedade da uva, a forma como as videiras são cultivadas e uma série de práticas empregadas na vinificação. Devido à sua composição, mostos e vinhos formam soluções ácido-básicas denominadas de soluções tampão, isto é, modificações na sua composição química não produzem variação significativa no pH. Assim são relativamente pequenas as variações no pH do mosto durante a fermentação alcoólica e maltoláctica (RIBERÉAU-GAYON *et al.*, 2006).

A maioria dos compostos do vinho tinto pertence à família dos polifenóis; esses existem em grande quantidade e podem ser classificados em grupos de flavonoides que se dividem em várias subclasses: flavan-3-ol (catequina, epicatequina etc.), flavonóis (quercetina, miricetina etc.) e os não-flavonoides que também são divididas em subclasses como, os ácidos fenólicos (cinâmico, benzoico etc.), aldeídos fenólicos (siraldeído etc.) e estilbenos (trans-resveratrol). Além disso, o vinho possui vitamínicos (riboflavina etc.) e aminoácidos (triptofano) (RODRÍGUEZ *et al.*, 2011; SAAD, 2016).

A cor do vinho tinto está associada à presença de um importante grupo de pigmentos fenólicos solúveis em água, chamadas de antocianinas, e sua coloração depende dos substituintes (OH, CH<sub>3</sub>) no C da aglicona. A cor do vinho depende também do pH, e varia de acordo com a copigmentação e interação da quelação dos íons com o etanol (CHEYNIER *et al.*, 1998; BOULTON, 2001).

As antocianinas em uvas vermelhas são os 3-glicosídeos de cinco agliconas: cianidina, petunidina, peonidina, delphinidina e malvidina. Esses compostos estão presentes como formas glicosiladas que podem ser acilados na porção de glicose por acetil, grupos cafeoil ou coumaroil. A malvidina 3-O-glicosídeo e seus derivados são as antocianinas mais representativas em uvas e vinhos, desempenhando papel essencial na qualidade sensorial do vinho (sabor, cor, adstringência). As antocianinas são extraídas do vermelho das peles de uva e, em alguns casos, de polpa de uva durante vinificação e estão presentes em quantidades relativamente grandes em vinhos tintos jovens. Estes pigmentos são gradualmente transformados por outros pigmentos derivados da antocianina durante armazenamento e envelhecimento do vinho, contribuindo para mudança da cor vermelho-

púrpura do vinho tinto jovem para uma cor mais vermelho-laranja de vinho tinto envelhecido. As transformações químicas de antocianinas formam novos pigmentos estáveis e sua relevância para a cor do vinho tem atraído muitos estudos em soluções modelo e em vinhos tintos (RIVAS *et al.*, 1995; FULCRAND *et al.*, 1996; MATEUS *et al.*, 2002; MATEUS, 2003).

A modificações de pH induzem variações de coloração. O pH abaixo de 2 faz com que as antocianinas apresentem intensa coloração avermelhada devido à forma de cátion flavílico (A<sup>+</sup>). Conforme o pH vai aumentando, as concentrações de A<sup>+</sup> e da base carbinol incolor (AOH) vão diminuindo. Com pH 4,5, as concentrações apresentam colorações de violeta ao azul e as formas de chalcona e quinona aumentam (GLORIES, 1984).

Estas variações de cores começaram a ser observadas no século XVII, quando Robert Boyle preparou um licor de violeta e percebeu que o extrato da flor se tornava vermelho em solução ácida e verde em meio básico. Hoje em dia, sabe-se que as antocianinas de flores, frutas e vegetais variam em função da sua acidez (coloração avermelhada), neutro (incolor) e alcalinidade (coloração esverdeada), e apresentam os pigmentos em função dos flavonoides (GROSS, 1987).

### 3.3.7 Vinho e Saúde

Segundo a história, o registro do vinho era muito utilizado pelos gregos. Hipócrates fez diversas observações sobre vinho e as suas propriedades medicinais, conforme constam em textos sobre a história da medicina (JOHSON, 2004).

Atualmente os vinhos tintos são reconhecidos em todo o mundo pelas suas propriedades benéficas para a saúde. Os compostos do vinho tinto atraem diferentes áreas de pesquisa; sendo assim, diversos estudos mostram que os vinhos tintos atuam na prevenção de diversas doenças, principalmente devido à sua composição fenólica (MAJKIĆ *et al.*, 2019).

Os compostos fenólicos estão presentes em frutas, vegetais, chás e vinhos e possuem diversos benefícios à saúde. Estudos epidemiológicos, clínicos e *in vitro*, mostram múltiplos efeitos biológicos relacionados aos compostos fenólicos da dieta, tais como: atividade antioxidante, anti-inflamatória, antimicrobiana e anti-carcinogênica

## REVISÃO BIBLIOGRÁFICA

---

(DELMAS *et al.*, 2001; CANTOS, ESPÍN & TOMÁS-BARBERÁN, 2002; BEER *et al.*, 2003; JANNIN & LATRUFFE, 2005).

Diversos estudos confirmam essas atividades e mostram o uso de vinhos tintos na prevenção de doenças cardiovasculares, dos sistemas urinário e respiratório, câncer entre outras (SHAN, CUIRONG & YUANJIANG, 2008; MAJKIĆ *et al.*, 2019; MRKUS, 2019;). Os polifenóis do vinho tinto, constituem vários antioxidantes poderosos, como flavonoides e estilbenos, e têm sido implicados em prevenção do câncer. Além disso, eles promovem a saúde humana sem efeitos colaterais reconhecíveis (MRKUS, 2019).







## 4.1 Apresentação

Neste capítulo serão apresentados os resultados da análise dos três métodos de seleção de variáveis empregados: ACO, GA e SW que foram aplicados para seleção dos sinais de fluorescência, visando a classificação geográfica das amostras de vinhos tintos das regiões da Campanha e da Serra Gaúcha, empregadas nesse trabalho.

### 4.1.1 Introdução

Em algumas técnicas analíticas muitas aplicações da calibração multivariada envolvem um enorme número de variáveis. Na grande maioria das vezes, o uso somente de algumas variáveis que contêm as informações relevantes, podem ocasionar maior segurança e praticidade na interpretação do modelo desenvolvido (FILHO & POPPI, 1999).

No processo de seleção de variáveis espectrais, foram empregados três métodos neste estudo: ACO, GA e SW, utilizando 73 amostras separadas em 2 Classes, sendo a Classe 1 para as amostras da Campanha e Classe 2 referentes as amostras da Serra Gaúcha. Essas amostras contemplam cerca de 20 pequenas regiões produtoras de vinhos.

O sinal de fluorescência registrado corresponde as matrizes de emissão (51 variáveis) de excitação (12 variáveis) (EEM) registradas em diferentes pHs. Cada matriz de fluorescência 2D foi desdobrada para obter um sinal de 1ª ordem.

Desse modo, foram selecionados os dados para apenas três estudos, empregando apenas dois pH analisados: pH 3, pH 7 e a fusão entre esses pHs (pH 3 + pH 7), que foram explorados por meio dos três métodos de seleção de variáveis, visando obter e avaliar modelos de classificação por análise discriminante linear (LDA) para determinação da origem geográfica dos vinhos tintos brasileiros produzidos nas duas regiões do Rio Grande do Sul.

### 4.1.2 Materiais e Métodos

#### 4.1.2.1 Solução Tampão e Produtos Químicos

As soluções tampão foram preparadas a partir dos valores de pH 3 e pH 7, utilizando os seguintes reagentes: Ácido fosfórico ( $H_3PO_4$ ) da Nuclear, Diadema, SP, Brasil; fosfato de potássio bibásico anidro ( $K_2HPO_4$ ). Água ultrapura (resistividade  $> 18M\Omega$  MilliQ, Millipore) foi usada para preparações das soluções.

#### 4.1.2.2 Amostras de Vinhos

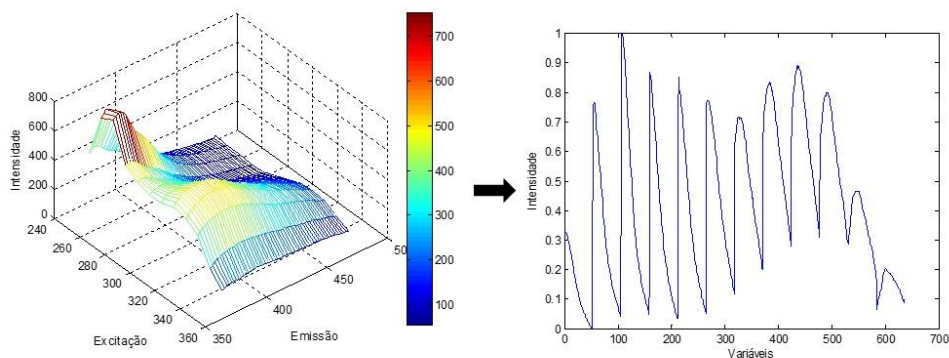
Foram utilizadas 73 amostras, sendo 53 procedentes da Serra Gaúcha e 20 da região da Campanha do Estado do Rio Grande do Sul, referentes a safras de 2017 e 2018, cedidas pela Embrapa Uva e Vinho – Bento Gonçalves. As amostras incluíram dez variedades de uvas tintas: Cabernet Sauvignon (CS), Pinot Noir (PN), Merlot (MT), Marselan (MN), Teroldego (TO), Tannat (TT), Malbec (MC), Cabernet Franc (CF), Ancellotta (AA) e Petit Syrah (PS). Para cada pH, as soluções tampão foram ajustadas usando um pHmetro de bancada (Ultra Denver Instrument<sup>®</sup>). Foram adicionados 450  $\mu$ L de amostra a aproximadamente 15 mL de tampão, para cada pH.

#### 4.1.2.3 Método Analítico

As medidas espectrais de fluorescência foram realizadas usando um espectrofotômetro Varian Cary Eclipse, com dois monocromadores Czerny-Turner e uma lâmpada de flash de xenônio conectada a um microcomputador de PC via interface serial RS232 (GPIB). As larguras das fendas dos monocromadores de excitação e emissão foram ambas definidas para 10 nm. Um detector PMT 600 V com software Cary Eclipse foi usado para aquisição dos dados. As medições EEM foram registradas como um conjunto de espectros de excitação de 244 a 340 nm e comprimentos de onda de emissão de 368 a 468 nm, totalizando 12 e 51 variáveis, respectivamente.

#### 4.1.2.4 Processos Quimiométricos

Para a obtenção dos dados de 1ª ordem, cada matriz EEM, foi desdobrada combinando os 12 espectros para cada comprimento de excitação, resultando num vetor com dados de fluorescência, conforme ilustrado na **Figura 12**. Os sinais resultantes foram normalizados entre 0-1, para cada amostra.



**Figura 12:** Esquema de transformação dos dados espectrais realizado.

Todos os cálculos quimiométricos foram realizados em ambiente Matlab® versão 7.1.0 (The MathWorks Inc., MA, EUA). Para a construção dos modelos de classificação (LDA) foi utilizado o método *leave on out*, as amostras de treinamento (2/3) e teste (1/3) foram selecionadas pelo método Kennard-Stone. Para a seleção das variáveis foi empregado, no Matlab, o pacote LDAVS\_gui (*linear discriminant analysis-variable selection toolbox*).

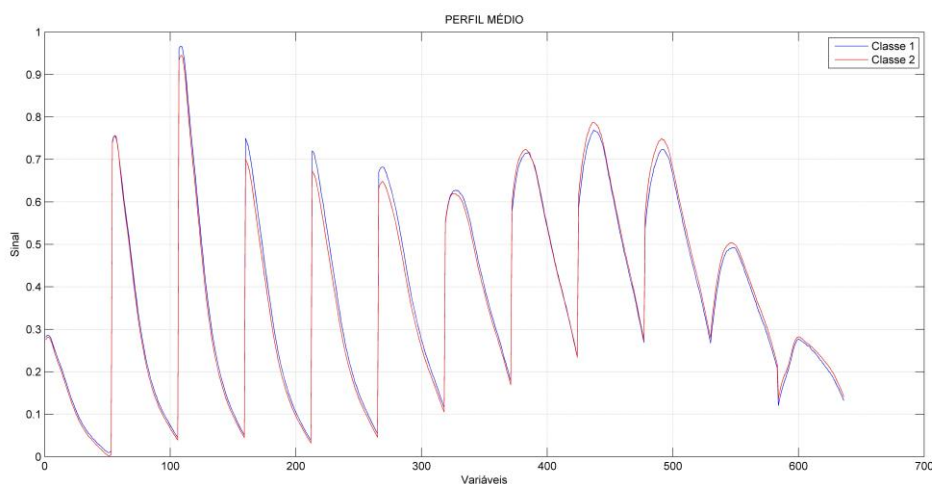
Foram empregados os seguintes parâmetros de entrada para o ACO: 100 formigas, 500 ciclos, 65% de taxa de evaporação do feromônio e *blind ants* de 0,35. Já para o GA foram empregados: 100 indivíduos, 1000 ciclos, 60% de probabilidade de cruzamento e 5% de probabilidade de mutação. Para ambos os algoritmos foi utilizado o valor 15 para o máximo de variáveis selecionadas. Para o SW o coeficiente de correlações múltiplas variou de 0,90 até 0,99.

### 4.1.3 Resultados e Discussão

#### 4.1.3.1 Resultados das Aplicações dos Algoritmos

##### 4.1.3.1.1 Resultados da Seleção do pH3

Inicialmente são apresentados na **Figura 13** os perfis médios das amostras de treinamento para as duas classes estudadas para o pH3, sendo a classe 1 referente às amostras da Campanha e classe 2 referentes as amostras da Serra Gaúcha. É possível observar sutis diferenças entre os sinais de fluorescência apresentados, sendo que na primeira faixa dos sinais desdobrados a classe da Campanha apresenta maior sinal, enquanto na segunda faixa a classe da Serra Gaúcha. Essas pequenas diferenças podem ser úteis na classificação de origem geográfica das amostras de vinho estudadas.



**Figura 13:** Gráficos das médias dos sinais espectrais das amostras de treinamento de vinhos tintos, para o pH3, sendo a Classe 1 referente as amostras da Campanha e Classe 2 referente as amostras da Serra Gaúcha.

Nas **Tabelas 1, 2, 3, 4, 5 e 6** são apresentados os resultados de classificação para os três métodos de seleção de variáveis empregados (ACO, GA e SW) para os dados de amostra de vinho ajustado para o pH3. Para os algoritmos heurísticos (ACO e GA) foram calculados também os valores médios das 5 realizações dos mesmos, buscando avaliar se as soluções apresentadas têm desempenho semelhante, além de comparar o desempenho das próprias heurísticas utilizadas.

Quando comparamos os resultados da tabela dos modelos ACO de treinamento (**Tabela 1**) com a de teste (**Tabela 2**), observamos que as figuras de mérito, em especial a taxa de precisão, apresentam melhores valores para as amostras de treinamento. O mesmo acontece com os modelos do GA (**Tabelas 3 e 4**).

**Tabela 1:** Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de treinamento para o pH3.

<i>Parâmetros</i>	<i>ACO1</i>	<i>ACO2</i>	<i>ACO3</i>	<i>ACO4</i>	<i>ACO5</i>	<i>Média</i>
Taxa de acerto (%)	90,3	94,1	82,6	92,6	86,4	89,2
Taxa de erro na classificação (%)	9,7	5,9	17,4	7,4	13,6	10,8
Nº de amostras classificadas erradas	5,0	4,0	7,0	5,0	6,0	5,4
<b>Classe 1</b>						
Precisão	0,750	0,765	0,714	0,722	0,733	0,737
Sensibilidade	0,923	1,000	0,769	1,000	0,846	0,908
Especificidade	0,882	0,882	0,882	0,853	0,882	0,876
<b>Classe 2</b>						
Precisão	0,968	1,000	0,909	1,000	0,938	0,963
Sensibilidade	0,882	0,882	0,882	0,853	0,882	0,876
Especificidade	0,923	1,000	0,769	1,000	0,846	0,908

**Tabela 2:** Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de teste para o pH3.

<i>Parâmetros</i>	<i>ACO1</i>	<i>ACO2</i>	<i>ACO3</i>	<i>ACO4</i>	<i>ACO5</i>	<i>Média</i>
Taxa de acerto (%)	49,2	56,0	74,6	77,4	44,8	60,4
Taxa de erro na classificação (%)	50,8	44,0	25,4	22,6	55,2	39,6
Nº de amostras classificadas erradas	12,0	8,0	6,0	5,0	12,0	8,6
<b>Classe 1</b>						
Precisão	0,273	0,400	0,556	0,625	0,222	0,415
Sensibilidade	0,429	0,286	0,714	0,714	0,286	0,486
Especificidade	0,556	0,833	0,778	0,833	0,611	0,722
<b>Classe 2</b>						
Precisão	0,714	0,750	0,875	0,882	0,688	0,782
Sensibilidade	0,556	0,833	0,778	0,833	0,611	0,722
Especificidade	0,429	0,286	0,714	0,714	0,286	0,486

## RESULTADOS

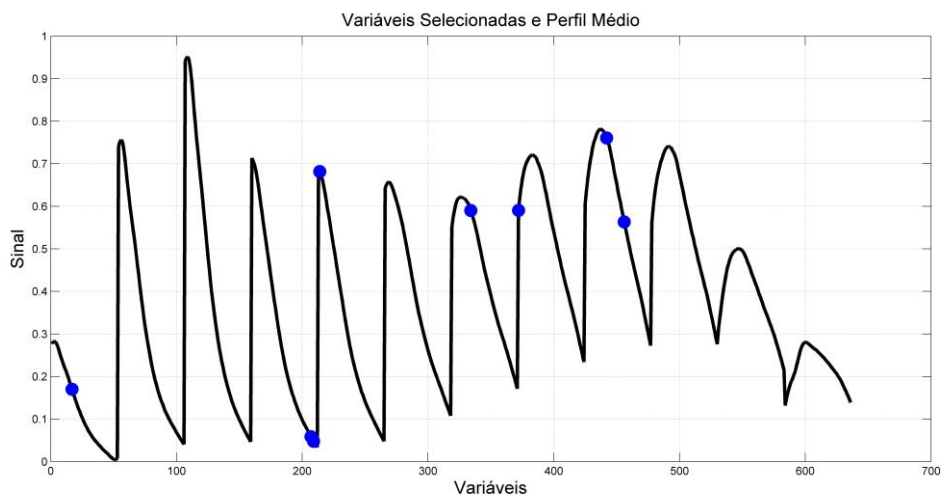
**Tabela 3:** Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de treinamento para o pH3.

<i>Parâmetros</i>	<i>GA1</i>	<i>GA2</i>	<i>GA3</i>	<i>GA4</i>	<i>GA5</i>	<i>Média</i>
Taxa de acerto (%)	78,7	81,1	86,4	83,5	90,3	84,0
Taxa de erro na classificação (%)	21,3	18,9	13,6	16,5	9,7	16,0
Nº de amostras classificadas erradas	8,0	8,0	6,0	8,0	5,0	7,0
<b>Classe 1</b>						
Precisão	0,692	0,667	0,733	0,647	0,750	0,698
Sensibilidade	0,692	0,769	0,846	0,846	0,923	0,815
Especificidade	0,882	0,853	0,882	0,824	0,882	0,865
<b>Classe 2</b>						
Precisão	0,882	0,906	0,938	0,933	0,968	0,925
Sensibilidade	0,882	0,853	0,882	0,824	0,882	0,865
Especificidade	0,692	0,769	0,846	0,846	0,923	0,815

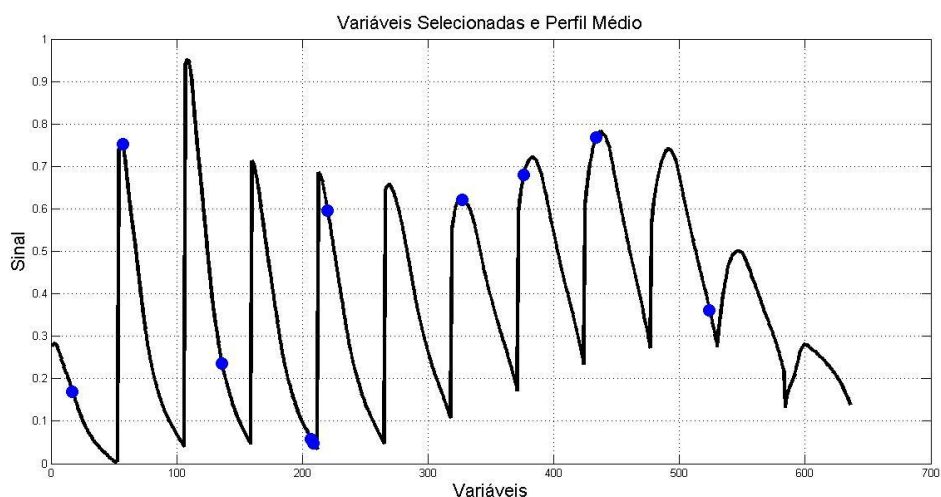
**Tabela 4:** Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de teste para o pH3.

<i>Parâmetros</i>	<i>GA1</i>	<i>GA2</i>	<i>GA3</i>	<i>GA4</i>	<i>GA5</i>	<i>Média</i>
Taxa de acerto (%)	67,5	70,2	65,9	73,0	84,5	72,2
Taxa de erro na classificação (%)	32,5	29,8	34,1	27,0	15,5	27,8
Nº de amostras classificadas erradas	7,0	6,0	6,0	5,0	4,0	5,6
<b>Classe 1</b>						
Precisão	0,500	0,571	0,600	0,667	0,667	0,601
Sensibilidade	0,571	0,571	0,429	0,571	0,857	0,600
Especificidade	0,778	0,833	0,889	0,889	0,833	0,844
<b>Classe 2</b>						
Precisão	0,824	0,833	0,800	0,842	0,938	0,847
Sensibilidade	0,778	0,833	0,889	0,889	0,833	0,844
Especificidade	0,571	0,571	0,429	0,571	0,857	0,600

Se considerarmos o desempenho dos modelos com base nas tabelas dos resultados para o conjunto teste (**Tabelas 2 e 4**), podemos selecionar os modelos ACO4 e GA5 que apresentam maior taxa de acerto (77,4% e 84,5%, respectivamente), sendo que o ACO4 selecionou 7 variáveis (**Figura 14**), enquanto o GA5 selecionou 9 variáveis (**Figura 15**).

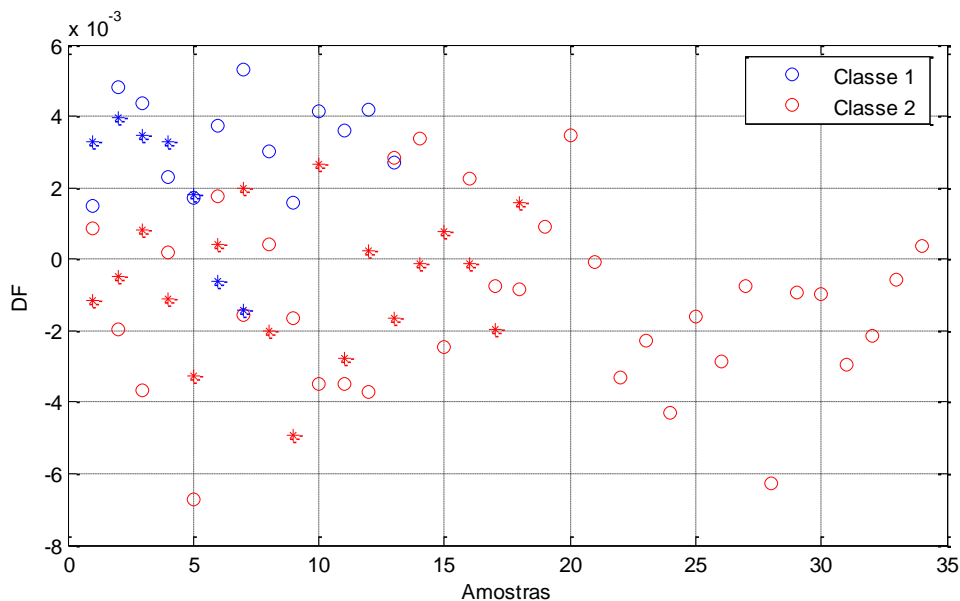


**Figura 14:** Gráfico das variáveis selecionadas para o modelo ACO4 e perfil médio dos espectros das amostras do pH3.

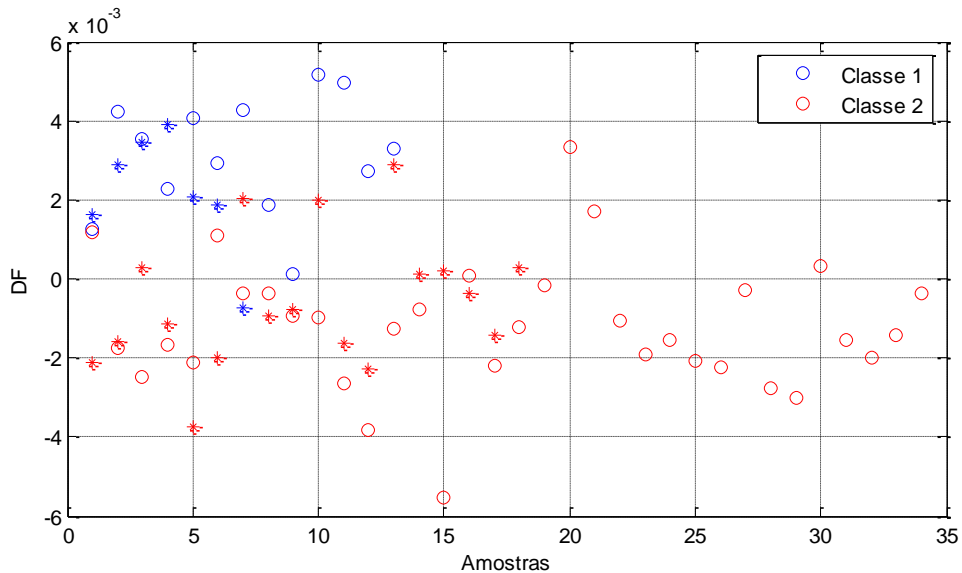


**Figura 15:** Gráfico das variáveis selecionadas para o modelo GA5 e perfil médio dos espectros das amostras do pH3.

Como consequência das variáveis selecionadas, os resultados das funções discriminantes podem ser observados nas **Figuras 16 e 17**, para os modelos ACO4 e GA5, respectivamente. Com base nas Figuras podemos considerar que o modelo GA5 apresentou capacidade discriminante superior ao ACO4, pois existe uma melhor separação das classes tanto dos conjuntos de treinamento, quanto de teste.



**Figura 16:** Gráfico dos resultados da LDA/ACO4 para as amostras do vinho tinto com pH3 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).



**Figura 17:** Gráfico dos resultados da LDA/GA5 para as amostras do vinho tinto com pH3 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).



Por outro lado, quando comparamos os modelos SW para diferentes valores de correlação para os resultados do conjunto teste para o pH3 (**Tabela 6**), poderíamos inicialmente supor que o modelo SW1 seria o mais adequado. Porém, de forma diferente que nos resultados dos algoritmos ACO e GA, os resultados para o conjunto de treinamento dos modelos SW variam muito, com total de amostras classificadas de forma equivocada variando de 5 até 12, sendo que o maior erro está associado ao modelo SW1. Já o modelo SW5, apesar de apresentar melhor desempenho para o conjunto de treinamento, com taxa de precisão de maior que 90%, não é adequado pois evidencia sobreajuste na modelagem, com uma queda na taxa de precisão para 71,8%, quando considerando os resultados do conjunto teste. Em virtude disso, o modelo SW3 foi selecionado como mais adequado para os dados de pH3, selecionando um total de 13 variáveis (**Figura 18**) e apresentado uma equivalente capacidade discriminatória como nos modelos heurísticos (ACO4 e GA5), conforme apresentado na **Figura 19**.

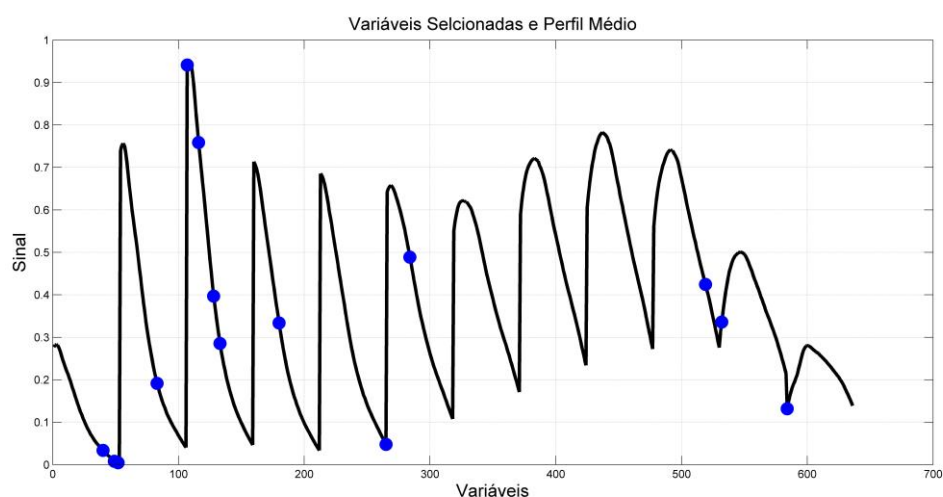
**Tabela 5:** Resultados da aplicação da formulação *Stepwise* (SW) para os dados espectrais das amostras de treinamento para o pH3.

<i>Parâmetros</i>	<i>SW1</i>	<i>SW2</i>	<i>SW3</i>	<i>SW4</i>	<i>SW5</i>
Coefficiente de correlações múltiplas	0,90	0,95	0,97	0,98	0,99
Taxa de acerto (%)	75,2	81,1	77,3	82,6	90,3
Taxa de erro na classificação (%)	24,8	18,9	22,7	17,4	9,7
Nº de amostras classificadas erradas	12,0	8,0	9,0	7,0	5,0
<i>Classe 1</i>					
Precisão	0,526	0,667	0,643	0,714	0,750
Sensibilidade	0,769	0,769	0,692	0,769	0,923
Especificidade	0,735	0,853	0,853	0,882	0,882
<i>Classe 2</i>					
Precisão	0,893	0,906	0,879	0,909	0,968
Sensibilidade	0,735	0,853	0,853	0,882	0,882
Especificidade	0,769	0,769	0,692	0,769	0,923

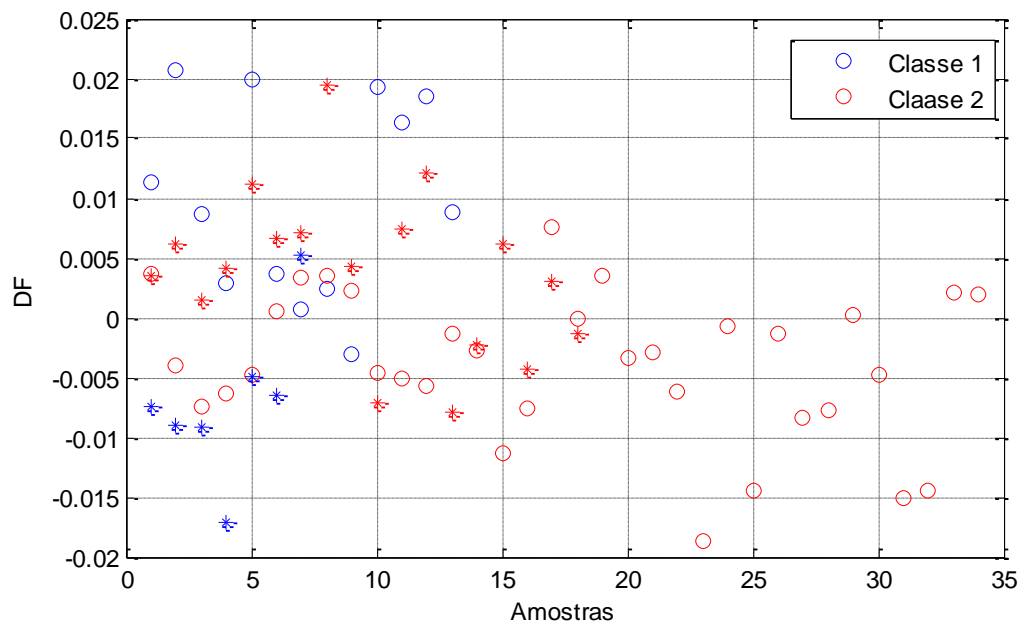
## RESULTADOS

**Tabela 6:** Resultados da aplicação da formulação **Stepwise (SW)** para os dados espectrais das amostras de teste para o pH3.

<i>Parâmetros</i>	<i>SW1</i>	<i>SW2</i>	<i>SW3</i>	<i>SW4</i>	<i>SW5</i>
Coeficiente de correlações múltiplas	0,90	0,95	0,97	0,98	0,99
Taxa de acerto (%)	94,4	81,8	84,5	70,2	71,8
Taxa de erro na classificação (%)	5,56	18,2	15,5	29,8	28,2
Nº de amostras classificadas erradas	2,0	5,0	4,0	6,0	7,0
<i>Classe 1</i>					
Precisão	0,778	0,600	0,667	0,571	0,500
Sensibilidade	1,000	0,857	0,857	0,571	0,714
Especificidade	0,889	0,778	0,833	0,833	0,722
<i>Classe 2</i>					
Precisão	1,000	0,933	0,938	0,833	0,867
Sensibilidade	0,889	0,778	0,833	0,833	0,722
Especificidade	1,000	0,857	0,857	0,571	0,714



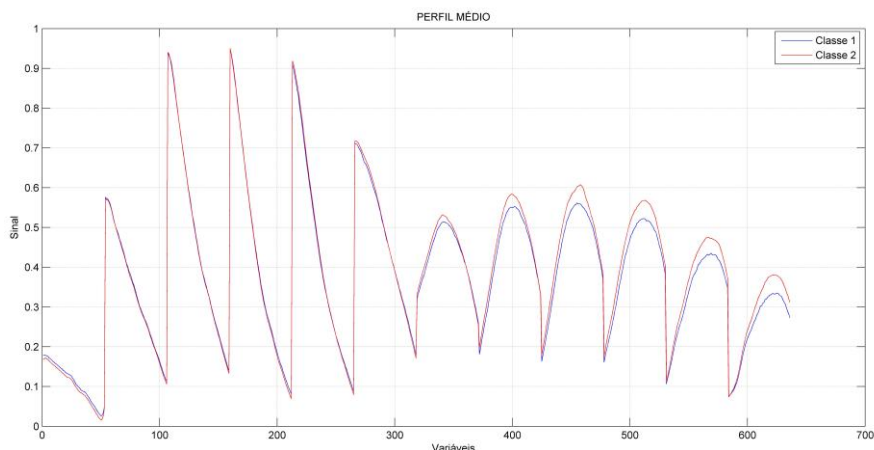
**Figura 18:** Gráfico das variáveis selecionadas para o modelo SW3 e perfil médio dos espectros das amostras do pH3.



**Figura 19:** Gráfico dos resultados da LDA/SW3 para as amostras do vinho tinto com pH3 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelho (teste).

#### 4.1.3.1.2. Resultados da Seleção do pH7

Os perfis médios das amostras de treinamento para o pH7 estão representados na **Figura 20**, onde as duas classes analisadas para esse pH foram a classe 1 referente as amostras da Campanha e classe 2 referentes as amostras da Serra Gaúcha. Podemos notar as diferenças entre os sinais de fluorescência apresentados, porém neste caso somente para a faixa final do sinal, onde a classe da Serra Gaúcha apresenta maior sinal.



**Figura 20:** Gráficos das médias dos sinais espectrais das amostras de treinamento de vinhos tintos, para o pH7, sendo a Classe 1 referente as amostras da Campanha e Classe 2 referente as amostras da Serra Gaúcha.

Nas **Tabelas 7, 8, 9, 10, 11 e 12** são apresentados os resultados de classificação para os três métodos de seleção de variáveis empregados (ACO, GA e SW) para os dados do pH7. Para os algoritmos heurísticos (ACO e GA) foram calculados também os valores médios das 5 realizações dos mesmos. Quando comparamos os resultados da tabela dos modelos ACO de treinamento (**Tabela 7**) com a de teste (**Tabela 8**), observamos que as figuras de mérito, em especial a taxa de precisão, também apresentam melhores valores para as amostras de treinamento. O mesmo acontece com os modelos do GA (**Tabelas 9 e 10**).

**Tabela 7:** Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de treinamento para o pH7.

<i>Parâmetros</i>	<i>ACO1</i>	<i>ACO2</i>	<i>ACO3</i>	<i>ACO4</i>	<i>ACO5</i>	<i>Média</i>
Taxa de acerto (%)	80,6	85,0	83,6	85,0	83,5	83,5
Taxa de erro na classificação (%)	19,5	15,0	16,5	15,0	16,5	16,5
Nº de amostras classificadas erradas	10,0	7,0	8,0	7,0	8,0	8,0
<b>Classe 1</b>						
Precisão	0,578	0,687	0,647	0,687	0,647	0,649
Sensibilidade	0,846	0,846	0,846	0,846	0,846	0,846
Especificidade	0,764	0,852	0,823	0,852	0,823	0,822
<b>Classe 2</b>						
Precisão	0,928	0,935	0,933	0,935	0,933	0,932
Sensibilidade	0,764	0,852	0,823	0,852	0,823	0,822
Especificidade	0,846	0,846	0,846	0,846	0,846	0,846

**Tabela 8:** Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de teste para o pH7.

<i>Parâmetros</i>	<i>ACO1</i>	<i>ACO2</i>	<i>ACO3</i>	<i>ACO4</i>	<i>ACO5</i>	<i>Média</i>
Taxa de acerto (%)	54,8	53,2	61,9	53,2	61,9	57,0
Taxa de erro na classificação (%)	45,2	46,8	38,1	46,8	38,1	43,0
Nº de amostras classificadas erradas	10,0	9,0	9,0	9,0	9,0	9,2
<b>Classe 1</b>						
Precisão	0,333	0,333	0,400	0,333	0,400	0,359
Sensibilidade	0,428	0,285	0,571	0,285	0,571	0,428
Especificidade	0,666	0,777	0,666	0,777	0,666	0,710
<b>Classe 2</b>						
Precisão	0,750	0,736	0,800	0,736	0,800	0,764
Sensibilidade	0,666	0,777	0,666	0,777	0,666	0,710
Especificidade	0,428	0,285	0,571	0,285	0,571	0,428

## RESULTADOS

**Tabela 9:** Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de treinamento para o pH7.

<i>Parâmetros</i>	<i>GA1</i>	<i>GA2</i>	<i>GA3</i>	<i>GA4</i>	<i>GA5</i>	<i>Média</i>
Taxa de acerto (%)	88,8	69,9	85,0	83,5	88,8	83,2
Taxa de erro na classificação (%)	12,0	30,1	15,0	16,5	11,2	17,0
Nº de amostras classificadas erradas	6,0	14,0	7,0	8,0	6,0	8,2
<b>Classe 1</b>						
Precisão	0,705	0,473	0,687	0,647	0,705	0,643
Sensibilidade	0,923	0,692	0,846	0,846	0,923	0,846
Especificidade	0,852	0,705	0,852	0,823	0,852	0,816
<b>Classe 2</b>						
Precisão	0,966	0,857	0,935	0,933	0,966	0,931
Sensibilidade	0,852	0,705	0,852	0,823	0,852	0,816
Especificidade	0,923	0,692	0,846	0,846	0,923	0,846

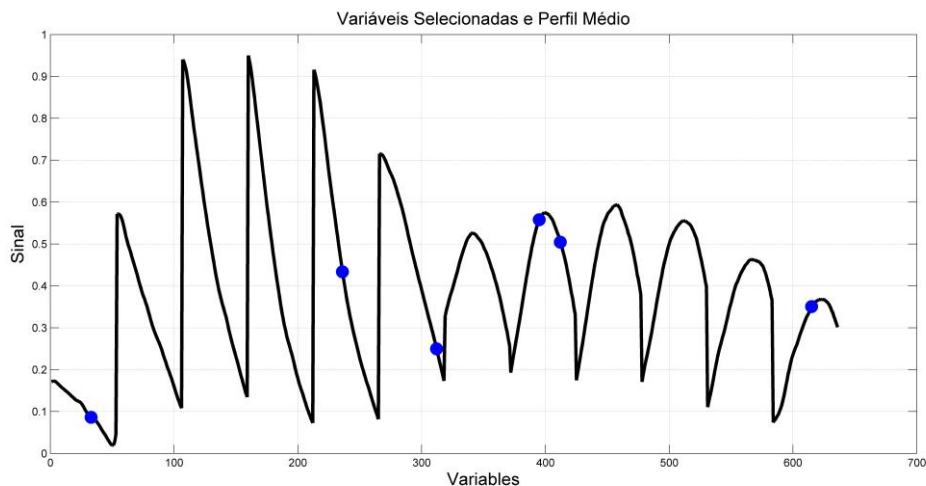
**Tabela 10:** Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de teste para o pH7.

<i>Parâmetros</i>	<i>GA1</i>	<i>GA2</i>	<i>GA3</i>	<i>GA4</i>	<i>GA5</i>	<i>Média</i>
Taxa de acerto (%)	43,7	39,3	30,0	64,7	54,8	46,5
Taxa de erro na classificação (%)	56,3	60,7	69,0	35,3	45,2	53,3
Nº de amostras classificadas erradas	14,0	14,0	17,0	8,0	10,0	12,6
<b>Classe 1</b>						
Precisão	0,230	0,181	0,142	0,444	0,333	0,266
Sensibilidade	0,428	0,285	0,285	0,571	0,428	0,399
Especificidade	0,444	0,500	0,333	0,722	0,666	0,533
<b>Classe 2</b>						
Precisão	0,666	0,642	0,545	0,812	0,750	0,683
Sensibilidade	0,444	0,500	0,333	0,722	0,666	0,533
Especificidade	0,428	0,285	0,285	0,571	0,428	0,399

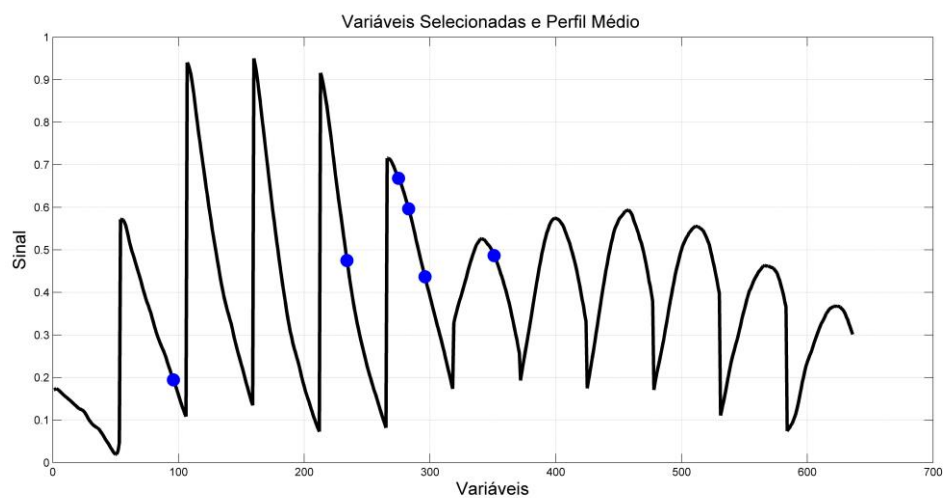
Se considerarmos o desempenho dos modelos com base nas tabelas dos resultados para o conjunto teste (**Tabelas 8 e 10**), podemos selecionar os modelos ACO5 e GA4 que apresentam maior taxa de precisão (61,9% e 64,7%, respectivamente), sendo que ambos modelos selecionaram 6 variáveis (**Figura 21 e 22**).

Cabe destacar que na **Tabela 8**, os modelos ACO3 e ACO5 apresentam resultados iguais, bem como na **Tabela 7**, que se refere aos modelos empregando os dados das amostras de treinamento. Ao observar as variáveis selecionadas para estes modelos,

constatamos que foram as mesmas 6 variáveis, logo representam uma única solução. Isto é raro, quando empregamos heurísticas para a seleção de variáveis para problemas desta natureza, mas pode acontecer principalmente para elevados valores de iteração.



**Figura 21:** Gráfico das variáveis selecionadas para o modelo ACO5 e perfil médio dos espectros das amostras do pH7.

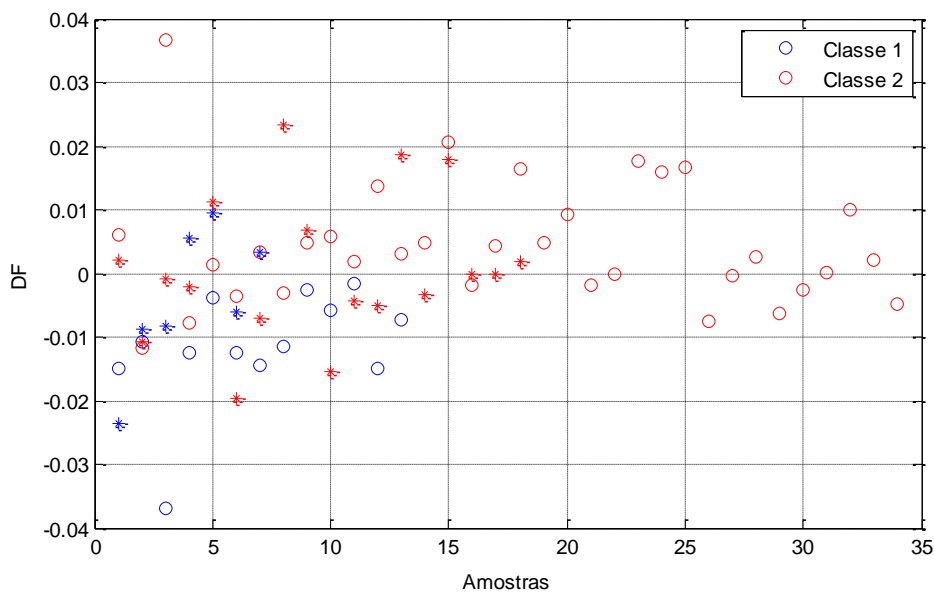


**Figura 22:** Gráfico das variáveis selecionadas para o modelo GA4 e perfil médio dos espectros das amostras do pH7.

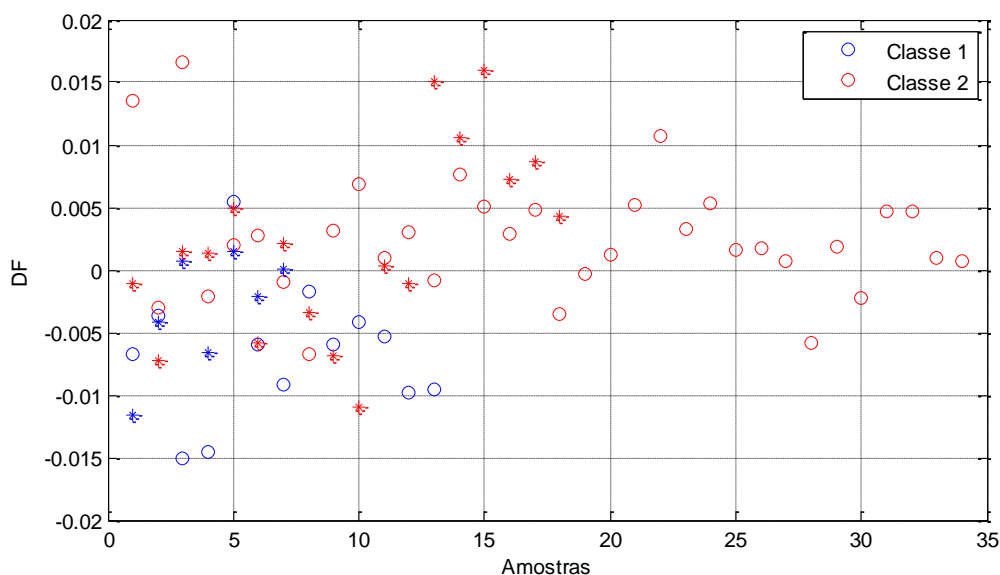
Como consequência das variáveis selecionadas, os resultados das funções discriminantes podem ser observados nas **Figuras 23 e 24**, para os modelos ACO5 e GA4, respectivamente. Com base nas Figuras podemos considerar que o model

## RESULTADOS

apresentou capacidade discriminante superior ao ACO4, pois existe uma melhor separação das classes tanto dos conjuntos de treinamento, quanto de teste.



**Figura 23:** Gráfico dos resultados da LDA/ACO5 para as amostras do vinho tinto com pH7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).



**Figura 24:** Gráfico dos resultados da LDA/GA4 para as amostras do vinho tinto com pH7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).



Ao compararmos os modelos SW, para diferentes valores de correlação para os resultados do conjunto teste para o pH7 (**Tabela 12**), observamos que os modelos SW1, SW3, SW4 e SW5 apresentam os mesmos valores para as figuras de méritos apresentadas, com taxa de acerto de 61,9%. Neste caso foi selecionado o modelo SW4, pois o mesmo apresentou melhor desempenho com relação as amostras de treinamento (**Tabela 11**). Para este modelo foram selecionadas 17 variáveis (**Figura 24**) e também uma equivalente capacidade discriminatória como nos modelos heurísticos (ACO5 e GA4), conforme apresentado na **Figura 25**. Deve-se salientar ainda que a maior parte das variáveis selecionadas por este modelo, estão na primeira faixa dos sinais desdobrados, o que pode justificar o baixo desempenho obtido para as amostras de teste.

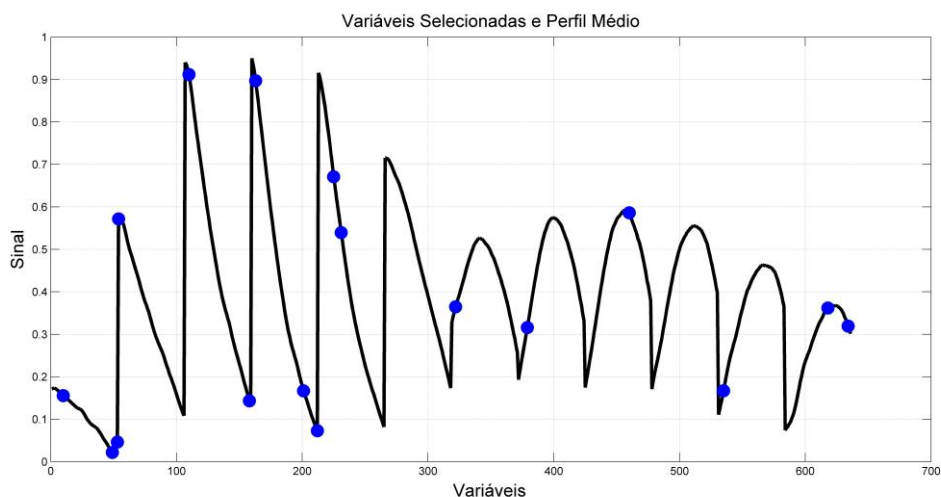
**Tabela 11:** Resultados da aplicação da formulação *Stepwise* (SW) para os dados espectrais das amostras de treinamento para o pH7.

<i>Parâmetros</i>	<i>SW1</i>	<i>SW2</i>	<i>SW3</i>	<i>SW4</i>	<i>SW5</i>
Coeficiente de correlações múltiplas	0,90	0,95	0,97	0,98	0,99
Taxa de acerto (%)	73,8	74,3	83,5	90,3	86,4
Taxa de erro na classificação (%)	26,2	25,7	16,5	9,7	13,6
Nº de amostras classificadas erradas	13,0	11,0	8,0	5,0	6,0
<i>Classe 1</i>					
Precisão	0,500	0,562	0,647	0,750	0,733
Sensibilidade	0,769	0,692	0,846	0,923	0,846
Especificidade	0,706	0,794	0,823	0,882	0,882
<i>Classe 2</i>					
Precisão	0,889	0,871	0,933	0,968	0,938
Sensibilidade	0,706	0,794	0,823	0,882	0,882
Especificidade	0,769	0,692	0,846	0,923	0,846

## RESULTADOS

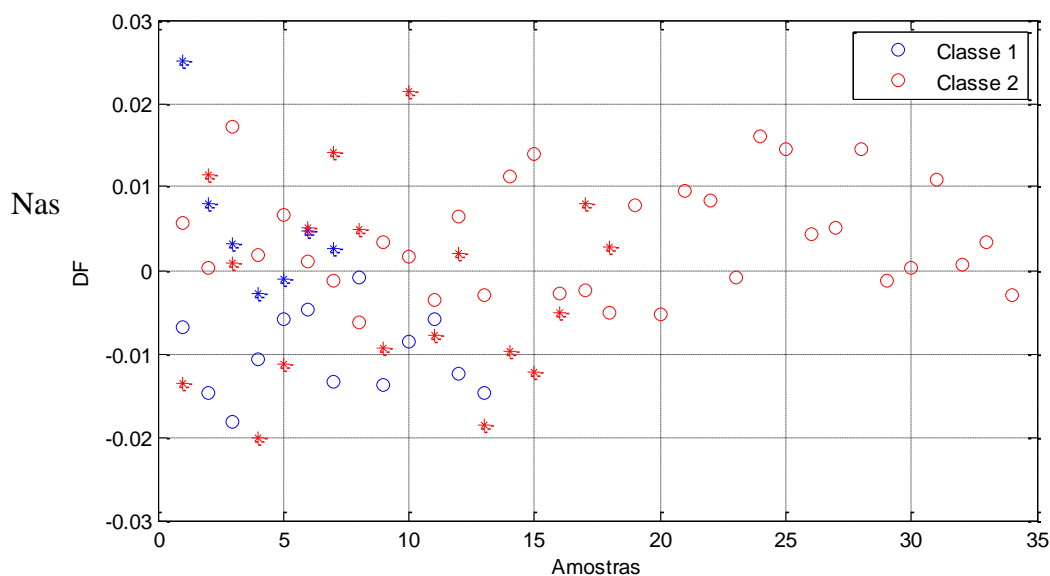
**Tabela 12:** Resultados da aplicação da formulação *Stepwise* (SW) para os dados espectrais das amostras de teste para o pH7.

<i>Parâmetros</i>	<i>SW1</i>	<i>SW2</i>	<i>SW3</i>	<i>SW4</i>	<i>SW5</i>
Coefficiente de correlações múltiplas	0,90	0,95	0,97	0,98	0,99
Taxa de acerto (%)	61,9	56,3	61,9	61,9	61,9
Taxa de erro na classificação (%)	38,1	43,7	38,1	38,1	38,1
Nº de amostras classificadas erradas	9,0	11,0	9,0	9,0	9,0
<i>Classe 1</i>					
Precisão	0,400	0,333	0,400	0,400	0,400
Sensibilidade	0,571	0,571	0,571	0,571	0,571
Especificidade	0,667	0,556	0,667	0,667	0,667
<i>Classe 2</i>					
Precisão	0,800	0,769	0,800	0,800	0,800
Sensibilidade	0,667	0,556	0,667	0,667	0,667
Especificidade	0,571	0,571	0,571	0,571	0,571



**Figura 25:** Gráfico das variáveis selecionadas para o modelo SW4 e perfil médio dos espectros das amostras do pH7.

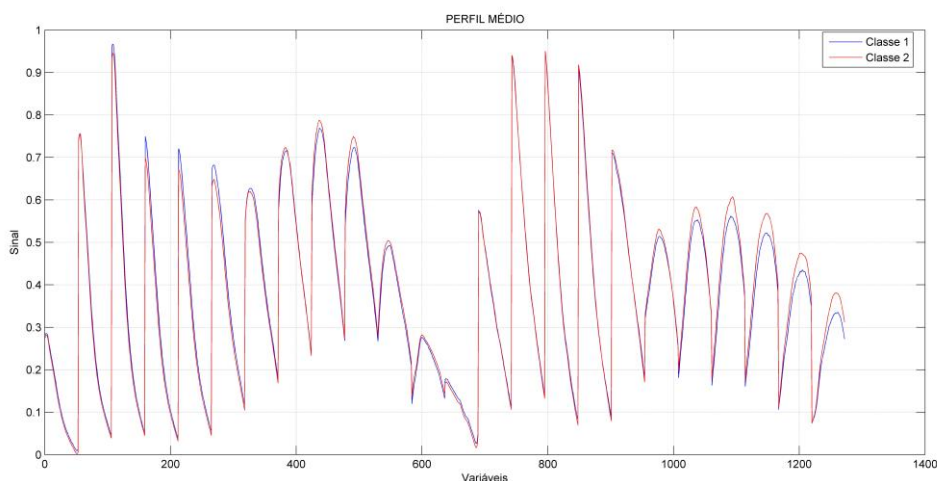
Para a fusão dos dados dos pH3 e pH7, mostrado na **Figura 28**, temos os perfis médios das amostras de treinamento para as duas classes estudadas, sendo a classe 1 referente as amostras da Campanha e classe 2 referentes as amostras da Serra Gaúcha. É possível observar as mesmas diferenças, já discutidas anteriormente.



**Figura 26:** Gráfico dos resultados da LDA/SW4 para as amostras do vinho tinto com pH7 das duas regiões, sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).

#### 4.1.3.1.3 Resultados da Seleção para a fusão dos pH3 e pH7

**Tabelas 13, 14, 15, 16, 17 e 18** são apresentados os resultados de classificação para os três métodos de seleção de variáveis empregados (ACO, GA e SW) para a fusão dos dados do pH3 e do pH7. Para os algoritmos heurísticos (ACO e GA) foram calculados também os valores médios das 5 realizações dos mesmos. Quando comparamos os resultados da tabela dos modelos ACO de treinamento (**Tabela 13**) com a de teste (**Tabela 14**), também observamos que as figuras de mérito apresentam melhores valores para as amostras de treinamento. O mesmo acontece com os modelos do GA (**Tabelas 15 e 16**).



**Figura 27:** Gráficos das médias dos sinais espectrais das amostras de treinamento de vinhos tintos, para a fusão dos dados do pH3 e do pH7, sendo a Classe 1 referente as amostras da Campanha e Classe 2 referente as amostras da Serra Gaúcha.

**Tabela 13:** Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de treinamento para a fusão dos dados do pH3 e do pH7.

<i>Parâmetros</i>	<i>ACO1</i>	<i>ACO2</i>	<i>ACO3</i>	<i>ACO4</i>	<i>ACO5</i>	<i>Média</i>
Taxa de acerto (%)	82,9	80,5	72,9	79,6	91,7	81,5
Taxa de erro na classificação (%)	17,1	19,5	27,1	20,4	8,3	18,5
Nº de amostras classificadas erradas	10,0	10,0	12,0	9,0	4,0	9,0
<b>Classe 1</b>						
Precisão	0,571	0,578	0,529	0,625	0,800	0,620
Sensibilidade	0,923	0,846	0,692	0,769	0,923	0,830
Especificidade	0,735	0,764	0,764	0,823	0,911	0,799
<b>Classe 2</b>						
Precisão	0,961	0,928	0,866	0,903	0,968	0,925
Sensibilidade	0,735	0,764	0,764	0,823	0,911	0,799
Especificidade	0,923	0,846	0,692	0,769	0,923	0,829

**Tabela 14:** Figuras de mérito para os modelos com aplicação do Algoritmo Colônia de Formigas (ACO) para os dados espectrais das amostras de teste para a fusão dos dados do pH3 e do pH7.

<i>Parâmetros</i>	<i>ACO1</i>	<i>ACO2</i>	<i>ACO3</i>	<i>ACO4</i>	<i>ACO5</i>	<i>Média</i>
Taxa de acerto (%)	36,5	60,3	57,5	73,5	66,3	58,8
Taxa de erro na classificação (%)	63,5	39,7	42,5	26,6	33,7	41,2
Nº de amostras classificadas erradas	15,0	8,0	9,0	8,0	9,0	9,8
<b>Classe 1</b>						
Precisão	0,166	0,428	0,375	0,461	0,416	0,369
Sensibilidade	0,285	0,428	0,428	0,857	0,714	0,542
Especificidade	0,444	0,777	0,722	0,611	0,611	0,510
<b>Classe 2</b>						
Precisão	0,615	0,777	0,764	0,916	0,846	0,783
Sensibilidade	0,444	0,777	0,722	0,611	0,611	0,633
Especificidade	0,285	0,428	0,428	0,857	0,714	0,542

**Tabela 15:** Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de treinamento para a fusão dos dados do pH3 e do pH7.

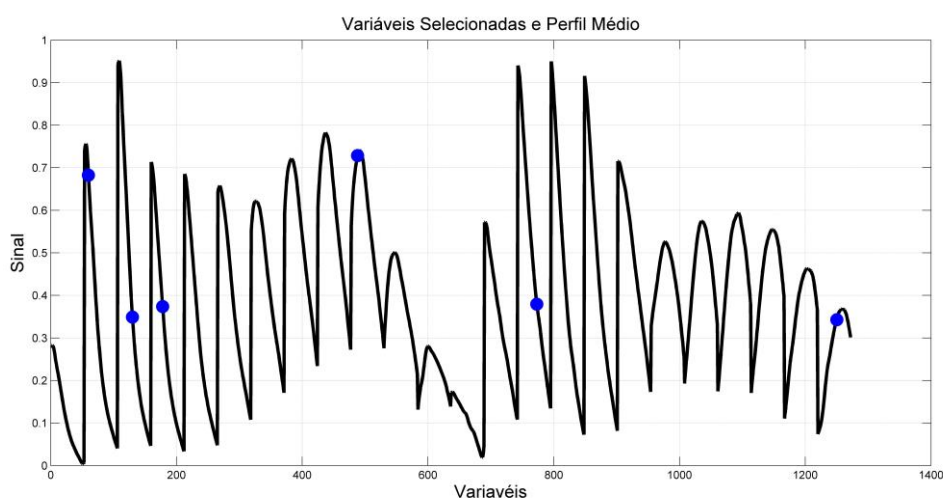
<i>Parâmetros</i>	<i>GAI</i>	<i>GA2</i>	<i>GA3</i>	<i>GA4</i>	<i>GA5</i>	<i>Média</i>
Taxa de acerto (%)	89,4	78,2	85,0	91,2	96,2	88,0
Taxa de erro na classificação (%)	10,6	21,8	15,0	8,8	3,8	12,0
Nº de amostras classificadas erradas	4,0	10,0	7,0	6,0	1,0	5,6
<b>Classe 1</b>						
Precisão	0,846	0,588	0,687	0,684	1,000	0,761
Sensibilidade	0,846	0,769	0,846	1,000	0,923	0,876
Especificidade	0,941	0,794	0,852	0,823	1,000	0,882
<b>Classe 2</b>						
Precisão	0,941	0,900	0,935	1,000	0,971	0,949
Sensibilidade	0,941	0,794	0,852	0,823	1,000	0,882
Especificidade	0,846	0,769	0,846	1,000	0,923	0,876

## RESULTADOS

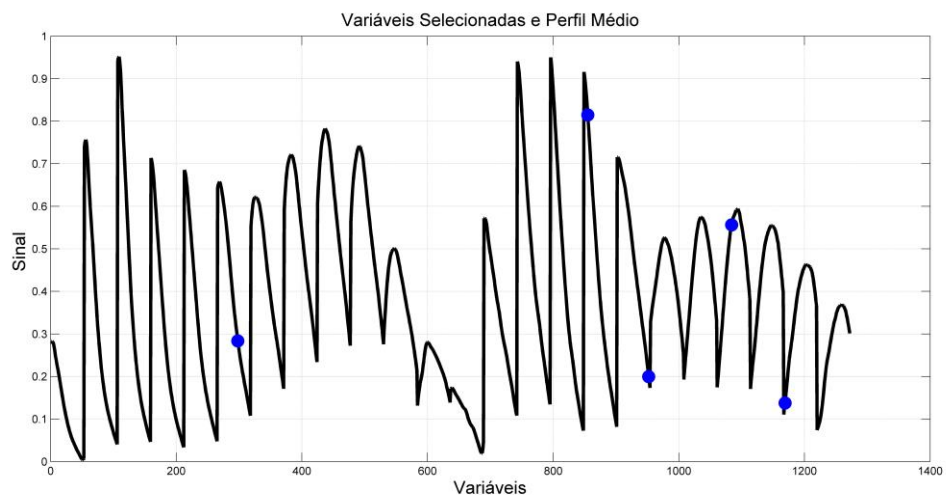
**Tabela 16:** Figuras de mérito para os modelos com aplicação do Algoritmo Genético (GA) para os dados espectrais das amostras de teste para a fusão dos dados do pH3 e do pH7.

<i>Parâmetros</i>	<i>GA1</i>	<i>GA2</i>	<i>GA3</i>	<i>GA4</i>	<i>GA5</i>	<i>Média</i>
Taxa de acerto (%)	54,8	77,4	61,9	67,5	59,1	64,1
Taxa de erro na classificação (%)	45,2	22,6	38,1	32,5	40,9	35,9
Nº de amostras classificadas erradas	10,0	5,0	9,0	7,0	10,0	8,2
<b>Classe 1</b>						
Precisão	0,333	0,625	0,400	0,500	0,363	0,498
Sensibilidade	0,428	0,714	0,571	0,571	0,571	0,571
Especificidade	0,666	0,833	0,666	0,777	0,611	0,710
<b>Classe 2</b>						
Precisão	0,750	0,882	0,800	0,823	0,785	0,808
Sensibilidade	0,666	0,833	0,666	0,777	0,611	0,710
Especificidade	0,428	0,714	0,571	0,571	0,571	0,571

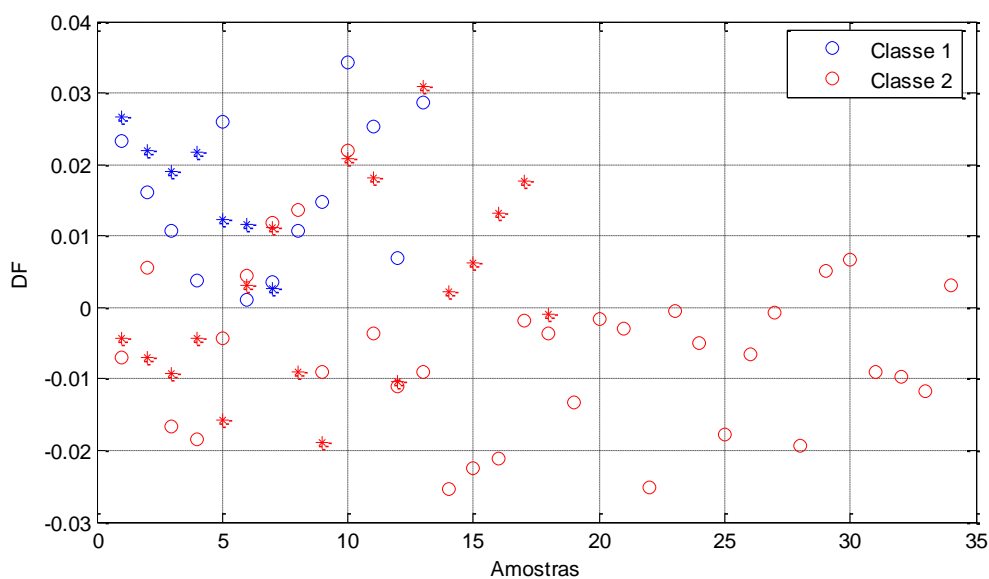
Se considerarmos o desempenho dos modelos com base nas tabelas dos resultados para o conjunto teste (**Tabelas 14 e 16**), podemos selecionar os modelos ACO4 e GA2 que apresentam maior taxa de precisão (73,5% e 77,4,5%, respectivamente), sendo que o ACO4 selecionou 6 variáveis (**Figura 27**), enquanto o GA2 selecionou 5 variáveis (**Figura 29**). Porém, ambos apresentam desempenho semelhante para as funções discriminantes resultantes (**Figuras 30 e 31**).



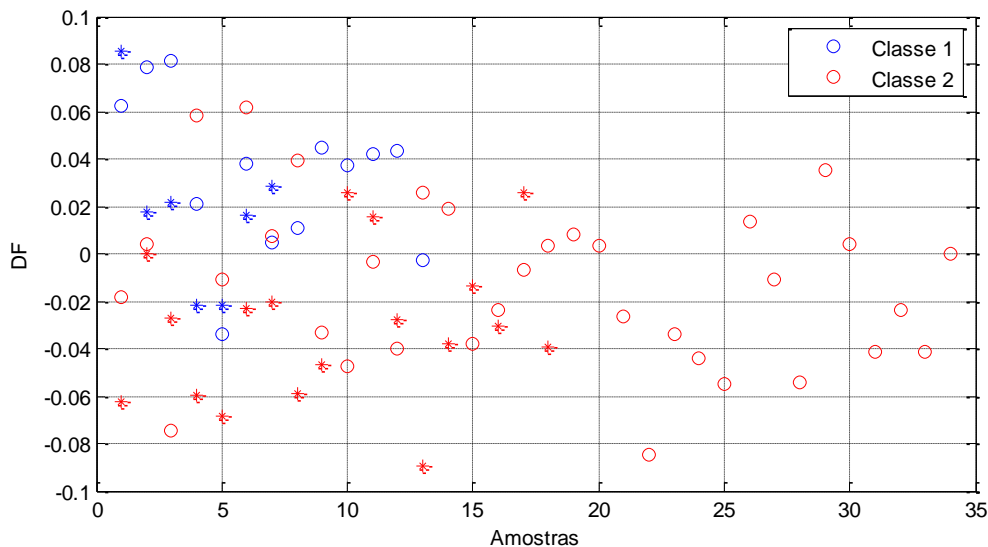
**Figura 28:** Gráfico das variáveis selecionadas para o modelo ACO4 e perfil médio dos espectros das amostras para a fusão dos dados do pH3 e do pH7.



**Figura 29:** Gráfico das variáveis selecionadas para o modelo GA2 e perfil médio dos espectros das amostras para a fusão dos dados do pH3 e do pH7.



**Figura 30:** Gráfico dos resultados da LDA/ACO4 para as amostras do vinho tinto com pH3 e pH7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).



**Figura 31:** Gráfico dos resultados da LDA/GA2 para as amostras do vinho tinto com pH3 e pH 7 das duas regiões. Sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).

Já quando comparamos os modelos SW para diferentes valores de correlação para os resultados do conjunto teste para a fusão dos dados do pH3 e do pH7 (**Tabela 18**), observamos que o modelo SW3 apresenta excelente taxa de acerto (>90%), além de apresentar os melhores valores para as demais figuras de mérito para ambas as classes. Além disso, o desempenho para os resultados para o conjunto de treinamento (**Tabela 17**) também foram comparáveis aos modelos com seleção heurística (ACO e GA).

**Tabela 17:** Resultados da aplicação da formulação *Stepwise* (SW) para os dados espectrais das amostras de treinamento para a fusão dos dados dos pH3 e pH7.

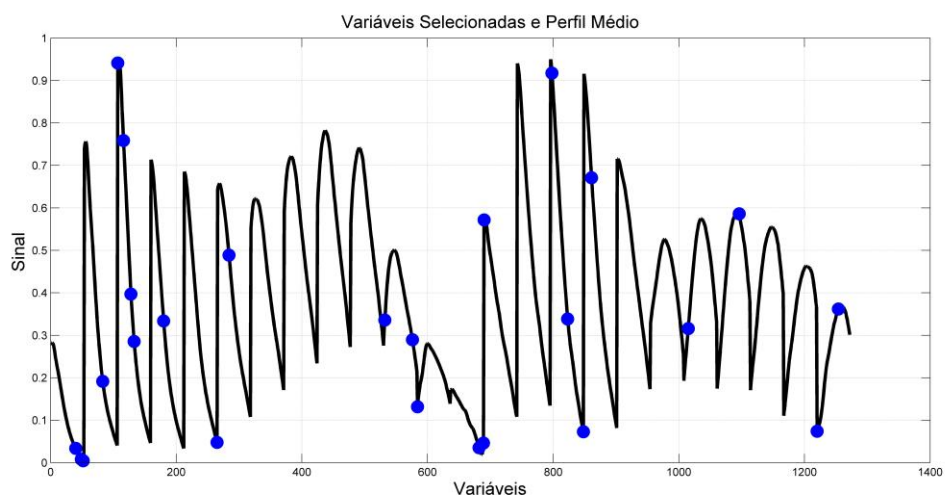
<i>Parâmetros</i>	<i>SW1</i>	<i>SW2</i>	<i>SW3</i>	<i>SW4</i>	<i>SW5</i>
Coefficiente de correlações múltiplas	0,90	0,95	0,97	0,98	0,99
Taxa de acerto (%)	85,9	94,1	93,2	93,2	98,5
Taxa de erro na classificação (%)	14,1	5,9	6,8	6,8	1,5
Nº de amostras classificadas erradas	8,0	4,0	3,0	3,0	1,0
<b><i>Classe 1</i></b>					
Precisão	0,631	0,764	0,857	0,857	0,929
Sensibilidade	0,923	1,000	0,923	0,923	1,000
Especificidade	0,794	0,882	0,941	0,941	0,971
<b><i>Classe 2</i></b>					
Precisão	0,964	1,000	0,970	0,970	1,000
Sensibilidade	0,794	0,882	0,941	0,941	0,971
Especificidade	0,923	1,000	0,923	0,923	1,000



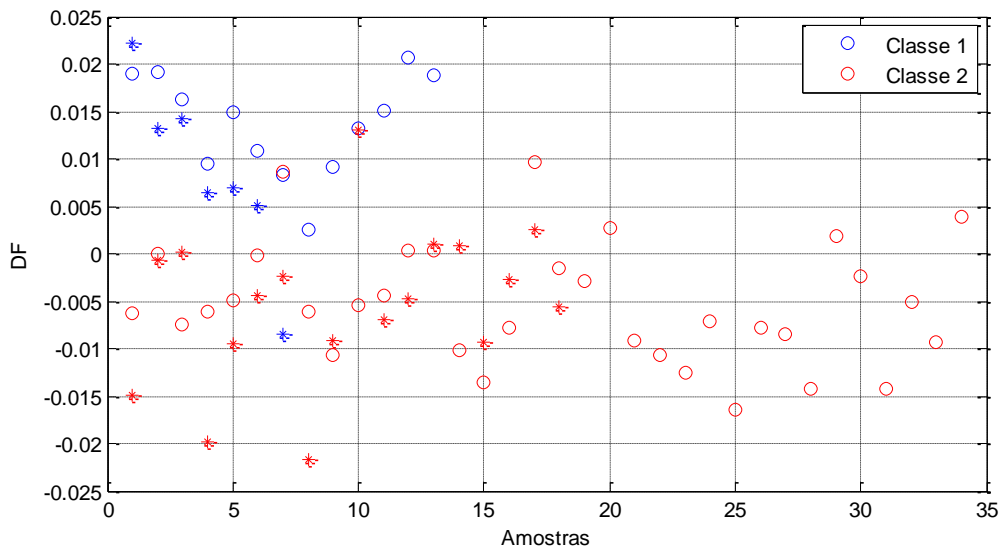
**Tabela 18:** Resultados da aplicação da formulação *Stepwise* (SW) para os dados espectrais das amostras de teste para a fusão dos dados dos pH3 e pH7.

<i>Parâmetros</i>	<i>SW1</i>	<i>SW2</i>	<i>SW3</i>	<i>SW4</i>	<i>SW5</i>
Coeficiente de correlações múltiplas	0,90	0,95	0,97	0,98	0,99
Taxa de acerto (%)	81,8	71,8	90,1	81,8	75,0
Taxa de erro na classificação (%)	18,2	28,2	9,9	18,2	25,0
Nº de amostras classificadas erradas	5,0	7,0	2,0	5,0	9,0
<i>Classe 1</i>					
Precisão	0,600	0,500	0,857	0,600	0,438
Sensibilidade	0,857	0,714	0,857	0,857	1,000
Especificidade	0,778	0,722	0,944	0,778	0,500
<i>Classe 2</i>					
Precisão	0,933	0,867	0,944	0,933	1,000
Sensibilidade	0,778	0,722	0,944	0,778	0,500
Especificidade	0,857	0,714	0,857	0,857	1,000

Para o modelo SW3, um total de 24 variáveis foram selecionadas, conforme ilustrado na **Figura 32**, contemplando tanto os dados do pH3, quanto do pH7. A partir do gráfico da **Figura 33**, verificamos que a função discriminante obtida foi capaz de fazer uma excelente separação das classes tanto para as amostras de treinamento, quanto para as de teste, independente da classe considerada.



**Figura 32:** Gráfico das variáveis selecionadas para o modelo SW3 e perfil médio dos espectros das amostras para a fusão dos dados dos do pH3 e do pH7.



**Figura 33:** Gráfico dos resultados da LDA/SW3 para as amostras do vinho tinto com pH3 e pH 7 das duas regiões, sendo a Classe 1 (Campanha) representada pelos círculos azuis (treinamento) e asteriscos azuis (teste) e para a Classe 2 (Serra Gaúcha) representada pelos círculos vermelhos (treinamento) e asteriscos vermelhos (teste).

Em resultados anteriores (com pH3 ou pH7), todos os modelos apresentaram melhor desempenho para a classe 2 (Serra Gaúcha), provavelmente devido ao fato dessa classe apresentar um número muito maior de amostras que a classe 1 (Campanha). Esse efeito é conhecido com desbalanceamento entre as classes, podendo afetar de forma significativa na construção dos modelos de classificação.

Porém, a fusão dos dados de fluorescência nos distintos pHs (pH3 + pH7) foi capaz de compensar o desbalanceamento, resultando num modelo com desempenho satisfatório tanto para as amostras de treinamento quanto teste, para ambas a classes estudadas. Este modelo, com certeza poderia auxiliar no reconhecimento da origem geográfica de amostras de vinho tinto provenientes das regiões da Campanha e Serra Gaúcha.





ABE, L. T., MOTA, R. V.; LAJOLO, F. M. & GENOVESE, M. I. Compostos Fenólicos e Capacidade Antioxidante de Cultivares de Uvas *Vitis labrusca L.* e *Vitis vinifera L.* **Ciências e Tecnologia de Alimentos**, v. 27, n. 2, p. 394-400, 2007.

ABE, Associação Brasileira de Enologia <<https://www.enologia.org.br/curiosidade/a-historia-do-vinho-no-brasil>>. Acesso em: 14 de dezembro de 2018.

ALLEGRI, 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. **Analytica Chimica Acta**, v. 699, p. 18-25, 2011.

ALVES, R. F.; NASCIMENTO, A. M. D. & NOGUERIA, J. M. F. Characterization of the Aroma Profile of Madeira Wine by Sorptive Extraction Techniques. **Analytica Chimica Acta**, v. 546, p. 11-21, 2011.

AZZOUZ, T. & TAULER R. Application of Multivariate Curve Resolution Alternating Least Squares (MCR-ALS) to the Quantitative Analysis of Pharmaceutical and Agricultural Samples. **Talanta**, v. 74, n. 5, p. 1201–10, 2008.

BALLABIO, D.; SKOV, T.; LEARDI, R. & BRO, R. Classification of GC-MS Measurements of Wines by Combining Data Dimension Reduction and Variable Selection Techniques. **Journal of Chemometrics**, v. 22, p. 457-463, 2008.

BANGALORE, A. S.; SHAFFER, R. E.; SMALL, G. W. & ARNOLD, M. A. Genetic Algorithm-Based Method for Selecting Wavelengths and Model Size for Use with Partial Least-Squares Regression: Application to Near-Infrared Spectroscopy. **Analytica Chimica Acta**, v. 68, n. 23, p. 4200-4212, 1996.

BEER, D.; JOUBERT, E.; GELDERBLUM, W. C. & MANLEY, M. Antioxidant activity of South African Red and White Cultivar Wines: Free Radical Scavenging. **Journal of Agricultural and Food Chemistry**, v. 51, p. 902-909, 2003.

BOOKSH, K. S. & KOWALSKI, B. R. Theory of Analytical Chemistry. **Analytical Chemistry**, v. 66, n. 15, p. 782-791A, 1994.

BOTELHO, R. V. & PIRES, E. J. P. **Viticultura como Opção de Desenvolvimento para os Campos Gerais**. Universidade Estadual de Ponta Grossa. Encontro de Fruticultura dos Campos Gerais, Campos Gerais. Anais. Ponta Grossa: 2009. v. 1, p. 40-54, 2009.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

BOULTON, R. The Copigmentation of Anthocyanins and Its Role in the Color of Red Wine: A Critical Review. **American Journal of Enology and Viticulture**, v. 52, n. 2, p. 67-87, 2001.

BOOKSH, K. S. & KOWALSKI, B. R. The Theory of Analytical Chemistry. **Analytical Chemistry**, v. 66, p. A782, 1994.

BRO, R. PARAFAC. Tutorial and Applications. **Chemometrics and Intelligent Laboratory Systems**. v. 38, p. 149-171, 1997.

BROWN, S. D., TAULER, R. & WALCZAK, B. **Comprehensive Chemometrics. Chemical and Biochemical Data Analysis**, Elsevier: Broomfield. 2009.

CANTOS, E.; ESPÍN, J. C. & TOMÁS-BARBERÁN, F. A. Varietal Differences Among the Polyphenol Profiles of Seven Table Grape Cultivars Studied by LC-DAD-MS-MS. **Journal of Agricultural and Food Chemistry**, v. 50, p. 5691-5696, 2002.

CANECA, A. R., PIMENTEL, M. F., GALVÃO, R. K., DA MATTA, C. E., DE CARVALHO, F. R., RAIMUNDO, I. M. J. R., PASQUINI, C. & ROHWEDDER, J. J. Assessment of Infrared Spectroscopy and Multivariate Techniques for Monitoring the Service Condition of Diesel-Engine Lubricating Oils. **Talanta**, v. 70, n. 2, p. 344-52, 2006.

CAO, L.; HARRINGTON, P. B. & LIU, J. SIMPLISMA and ALS Applied to Two-Way Nonlinear Wavelet Compressed Ion Mobility Spectra of Chemical Warfare Agent Simulants. **Analytical Chemistry**, v. 77, n. 8, p. 2575-2586, 2005.

CEBI, N.; DOGAN, C. E.; MESE, A. E.; OZDEMIR, D.; ARICI, M. & SAGDIC, O. A Rapid ATR-FTIR Spectroscopic Method for Classification of Gelatin Gummy Candies in Relation to the Gelatin Source. **Food Chemistry**, v. 277, p. 373-381, 2018.

CERSÓSIMO, R. & TEHEMAYER, C. A. **Vitivinicultura na Região da Campanha Gaúcha**. Anais do Salão Internacional de Ensino, Pesquisa e Extensão, v. 8, n. 4, 14 fev. 2020.

CUADROS-INOSTROZA, A.; GIAVALISCO, P.; HUMMEL, J.; ECKARDT, A.; WILLMITZER, L. & PÉÑA-CÓRTEZ, H. Discrimination of Wine Attributes by Metabolome Analysis. **Analytical Chemistry**, v. 82, n. 9, p. 3573-3580, 2010.

CHELOTTI, M. C. & PESSOA, V. L. S. **Assentamentos Rurais e as Transformações Territoriais no Pampa Gaúcho/RS/BR**. Bienal Coloquio de T Transformaciones Territoriales, 6, 2006, Santa Fe, Argentina. Santa Fe: UNL, 2006. Disponível em: <<http://bit.ly/18LyrCS>>. Acesso 13 de fevereiro de 2019.

CHEYNIER, V.; SOUQUET, J. M.; FULCRAND, H.; SARNI, P. & MOUTONET, M. **Stabilisation Tanins-anthocyanes, Données Générales**, Extrait de Les composés phénoliques, p. 1-9, 1998.

CHRISTIAN, G. D.; CALLIS, J. B. & DAVIDSON, E. R. **Array Detectors and Excitation– Emission Matrices in Multicomponent Analysis**. In: Wehry, E.L. **Modern Fluorescence Spectroscopy**. Plenum, New York, p. 111-165, 1981.

CLIMACO, P.; RICARDO-DA-SILVA, J.; LAUREANO, O. & TONIETTO, J. **O Clima Vitícola das Principais Regiões Produtoras de Uvas para Vinho de Portugal**. Universidade de Lisboa. Lisboa, p. 1-43, 2012.

DANZER, K. & CURRIE, L. A. Analytical Chemistry Division Commission on General Aspects of Analytical Chemistry. **Pure & Applied Chemistry**, v. 70, n. 4, p. 993-1014, 1998.

DARDEAU, R. **Vinho Fino Brasileiro**. Mauad X, Lapa: Rio de Janeiro, 2015.

DELMAS, D.; JANNIN, B. & LATRUFFE, N. Resveratrol: Preventing Properties Against Vascular Alterations and Ageing. **Molecular Nutrition & Food Research**, v. 49, p. 377-395, 2005.

DESTRAMPE, K. A. & HIEFTJE, G. M. New Instrumentation for Use in Excitation– Emission Fluorescence-polarization Measurements. **Applied Spectroscopy**, v. 47, n. 10, p. 1548-1554, 1993.

DESPAGNE, F. & MASSART, D. L. Neural Networks in Multivariate Calibration. **The Analyst Tutorial Review**, v. 123, p.157R-178R, 1998.

DIMITROVSKA, M.; TOMOVSKA, E. & BOCEVSKA, M. Characterization of Vranec, Cabernet Sauvignon and Merlot Wines Based on their Chromatic and Anthocyanin Profiles. **Journal of the Serbian Chemical Society**, v. 78, n. 9, p. 1309-1322, 2013.

DORIGO, M. & GAMBARDELLA, L. M. Ant Colonies for the Travelling Salesman Problem. **Biosystems**, v. 43, p. 73-81, 1997.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

DORIGO, M. & BLUM, C. Ant Colony Optimization Theory: A Survey. **Theoretical Computer Science**, v. 344, n. 2-3, p. 243-278, 2005.

DORIGO, M.; CARO, G. D. & GAMBARDELLA, L. M. Ant Algorithms for Discrete Optimization. **Artificial Life**, v. 5, n. 2, p.137-172, 1999.

DUDA, R.O., HART, P. E. & STORK, D. G. **Pattern Classification**, 2nd ed., John Wiley, New York, 2001.

ELSAYED, S.; EL-GOZAYER, K.; ALLAM, A. & SCHMIDHALTER, U. Passive Reflectance Sensing Using Regression and Multivariate Analysis to Estimate Biochemical Parameters of Different Fruits Kinds. **Scientia Horticulturae**, v. 243, p. 21-33, 2019.

EMBRAPA. Empresa Brasileiro de Pesquisa Agropecuário. **Artigo: Desempenho da vitivinicultura brasileira em 2015**. Disponível em: <<https://www.embrapa.br/busca-de-noticias/-/noticia/9952204/artigo-desempenho-da-vitivinicultura-brasileira-em-2015>>. Acesso em 13 de fevereiro de 2019.

FARIAS, R. Curso para Profissionais ciclo I. Rio de Janeiro: *Associação Brasileira de Sommeliers*, 2014.

FILHO, P. A. C. & POPPI, R. J. Algoritmo Genético em Química. **Química Nova**, v. 22, n. 3, 1999.

FILHO, R. C. D. M. **Aplicação de Métodos de Resolução de Curvas Multivariada a Dados Experimentais Gerados por Diferentes Técnicas Analíticas: Uma Visão Exploratória da Biossíntese da Violaceína**. Tese de Doutorado. Universidade Estadual de Campinas (UNICAMP). São Paulo. p. 97, 2011.

FERNANDES, D. D. de S.; ROMEO, F.; KREPPER, G.; DI NEZIO, M. S.; PISTONESI, M. F.; CENTURIÓN, M. E.; ARAÚJO, M. C. U. & DINIZ, P. H. G. D. Quantification and Identification of Adulteration in the Fat Content of Chicken Hamburgers Using Digital Images and Chemometric Tools. **LWT**, v. 100, p. 20-27, 2018.

FERREIRA, M. C.; ANTUNES, A. M., MELGO, M. S. & VOLPE, P. L. O. Quimiometria I: Calibração Multivariada, um Tutorial. **Química Nova**, v. 22, n. 5, p. 724-731. 1999.



FULCRAND, H.; CAMEIRA, S. P. J.; SARNI-MANCHADO, P.; CHEYNIER, V. & FAVRE, B. J. Structure of New Anthocyanin-derived Wine Pigments. **Jornal of Chemical Society Perkin**, v. 1, p. 735-739, 1996.

FLORES, S. S. **Desenvolvimento territorial sustentável a partir dos territórios do vinho: o caso dos “vinhos da campanha”**. Porto Alegre: UFRGS/ PPGEA, 2011. **Desenvolvimento Territorial Sustentável a Partir dos Territórios do Vinho: o Caso dos Vinhos da Campanha**. Disponível em: <<http://bit.ly/1H6A4d5>>. Acesso em 13 de fevereiro de 2019.

FLORES, S. S. & MEDEIROS, R. M. V. Ruralidades na Compreensão dos Territórios do Vinho e sua Identidade. **Campo-Território: Revista de Geografia Agrária**, v. 8, n. 15, p. 1-19, 2013.

FRANKE, B. M.; GREMAUD, G.; HADORN, R. & KREUZER, M. Geographic Origin of Meat – Elements of an Analytical Approach to its Authentication. **European Food Research and Technology**, v. 221, p. 493–503, 2005.

GASNIER, V. **O Livro do Vinho**. 2º ed. São Paulo: Publifolha. P. 352, 2015.

GALVÃO, R. K., ARAUJO, M. C., JOSÉ, G. E., PONTES, M. J., SILVA, E. C., SALDANHA, T. C. A Method for Calibration and Validation Subset Partitioning. **Talanta**, v. 67, n. 4, p. 736-40, 2005.

GEN, R. C. M. **Genetic Algorithms and Engineering Optimization**. John Wiley & Sons, Inc., 2000.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization, and Machine Learning**. Addison-Wesley, Reading 1989.

GÓMEZ, V. & CALLAO, M. P. Analytical Applications of Second-order Calibration Methods. **Analytica Chimica Acta**, v.627, p. 169-183, 2008.

GOSS, S.; ARON, S.; DENEUBOURG, J. L. & PASTEELS, J. Self-Organized Shortcuts in the Argentine Ant. **Naturwissenschaften**, v. 76, p. 579- 81, 1989.

GUERRA, C. C.; MANDELLI, F.; TONIETTO, J.; ZANUS, M. C.; CAMARGO, U. A. Conhecendo o Essencial Sobre Uvas e Vinhos. **Embrapa Uva e Vinho**. Bento Gonçalves. Documentos n° 48, p. 69, junho de 2009.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

GUERRA, C. C.; MANDELLI, F.; TONIETTO, J.; ZANUS, M. C. & CARAMGO, U. A. Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA). Embrapa Uva e Vinho. **Amigo do Vinho Brasileiro: Conhecendo o Essencial sobre Uvas e Vinhos**. Bento Gonçalves, Documentos n° 48, p. 67, janeiro de 2005.

GUERRA, C. C. Vinho tinto. In: VENTURINI FILHO, W.G. **Bebidas alcoólicas: Ciência e Tecnologia**. 1st ed. São Paulo: Blucher, Cap. 11. p. 209-233, 2010.

GHASEMI-VARNAMKHASTI, M. & FORINA, M. NIR Spectroscopy Coupled with Multivariate Computational Tools for Qualitative Characterization of the Aging of Beer, **Computers and Electronics in Agriculture**, v. 100, p. 34-40, 2014.

GLORIES, Y. La Couleur des Vins Rouges. II. Mesure, Origine et Interprétation, **Journal International des Sciences de la Vigne et du Vin**, v. 18, n. 4, p. 253-271, 1984.

GROSS, J.; **Pigments in Fruits**, Academic Press: London, 1987, p. 59.

HAUPT, S. E. & HAUPT, R. L. **Practical Genetic Algorithms**. John Wiley & Sons, 2004.

HEMMATEENEJAD, B.; SHAMSIPUR M.; ZARE-SHAHABADI V. & AKHOND, M. Building Optimal Regression Tree by Ant Colony System–Genetic Algorithm: Application to Modeling of Melting Points. **Analytica Chimica Acta**, v. 704, n. 1–2, p. 57-62, 2011.

HOLLAND, J. H. **Adaptation in Natural and Artificial Systems**. 1° ed, MIT Press, London, 1975.

IBGE, Instituto Brasileiro de Geografia e Estatística. **Levantamento Sistemático da Produção Agrícola (LSPA)**. Rio de Janeiro, v. 30, n. 1, p.1-81, janeiro, 2017.

IBRAVIN, Instituto Brasileiro de Vinho. **História do Vinho no Brasil**. 2018a. Disponível em: <<http://www.ibravin.org.br/Historia-do-Vinho-no-Brasil>>. Acesso em: 14 de dezembro de 2018.

IBRAVIN. Instituto Brasileiro de Vinho. **Qualidade Marca a Safra de Uva 2018 no Rio Grande do Sul**. 2018b. Disponível em: <<https://www.ibravin.org.br/Noticia/qualidade-marca-a-safra-de-uva-2018-no-rio-grande-do-sul/367>> . Acesso em 13 de fevereiro de 2019.

IUPAC, International Union of Pure and Applied Chemistry. Guidelines for Calibration in Analytical Chemistry: part 1- Fundamentals and Single Component Calibration. **Pure and Applied Chemistry**, v. 70, n. 4, p. 993-1014, 1998.

JACKSON, R. S. **Chemical Constituents of Grapes and Wine, in: Wine Science**, 3<sup>o</sup> ed., Academic Press, San Diego, p. 270-331, 2008a.

JACKSON, R. S. **Wine Science - Principles and Applications**. Chapter 10: Wine Laws, Authentication, and Geography. Academic Press, 2008, p. 77–640, 2008b.

JAUMONT, J.; de JUAN, A. & TAULER, R. MCR-ALS GUI 2.0: New Features and Applications. **Chemometrics and Intelligent Laboratory Systems**, v. 140, p. 1-12, 2015.

JIANG, J. H. & Ozaki, Y. Self-Modeling Curve Resolution (Smcr): Principles, Techniques, and Applications. **Applied Spectroscopy Reviews**, v. 37, n. 3, p. 321-45, 2002.

JOHSON, H. **The Story of Wine**. 2 Ed. Mitchell Beazley, Londres, p. 256, 2004.

JORGENSEN, S. E. & HALLING-SORENSEN, B. Drugs in the environment, **Chemosphere**, v. 40, n. 7, p. 691-699, 2000.

JUAN, A. de; RUTAN, S. C. & TAULER, R. Two-Way **Data Analysis: Multivariate Curve Resolution –Iterative Resolution Methods**. Capítulo: Comprehensive Chemometrics Chemical and Biochemical Data Analysis. **Elsevier** 2, p. 325–344, 2009.

JUAN, A. & TAULER, R. Multivariate Curve Resolution (MCR) from 2000: Progress in Concepts and Applications. **Critical Reviews in Analytical Chemistry**, v. 36, n.3-4, 2006.

JUAN, A.; JAUMOT, J. & TAULER, R. Multivariate Curve Resolution (MCR). Solving the Mixture Analysis Problem, **Analytical Methods**, v. 6, n. 14, p. 4964-4976, 2014.

JURADO, J. M.; ALCÁZAR, A.; PALACIOS-MORILLO, A. & PABLOS, F. Classification of Spanish DO White Wines According to their Elemental Profile by Means of Support Vector Machines. **Food Chemistry**, v. 135, p. 898-903, 2012.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

KUHN, G. B.; LOVATEL, J. L.; PREZOTTO, O. P.; RIVALDO, O. F.; MANDELLI, F.; SÔNEGO, O. R. **O Cultivo da Videira: Informações Básicas**. 2ª ed. Bento Gonçalves: EMBRAPA-CNPV, Circular Técnica 10, 1996. p. 60.

LAKOWICZ, J. R. **Principles of Fluorescence Spectroscopy**. 3º Ed., Springer, 2009.  
LILLA, C. **Introdução ao Mundo do Vinho**. Martins Fontes, São Paulo, p. 186, 2016.

LINDER, M. & SUNDBERG, R. Second-order calibration: Bilinear Least Squares Regression and a Simple Alternative. **Chemometrics and Intelligent Laboratory Systems**, v, 43, n. 1-2, 1998.

LINGARA, J. K; SRIKANTHA, N.; MOINUDDIN. K. & LOKESH K.S Swarm Intelligence Based Ant Colony Optimization (ACO) Approach for Maximizing the Lifetime of Heterogeneous Wireless Sensor Networks. **Journal of Engineering Research and Applications**, v. 3, n. 6, p. 167- 172, 2013.

LIU, J.; FAN, L.; LIU, Y.; MAO, L. & KAN, J. Application of Terahertz Spectroscopy and Chemometrics for Discrimination of Transgenic Camellia oil. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 206, p. 165-169, 2019.

LUYKX, D. M. A. M. & VAN R. S. M. An Overview of Analytical Methods for Determining the Geographical Origin of Food Products. **Food Chemistry**, v. 107, n. 2, p. 897-911, 2008.

MAJKIĆ, T. M.; TOROVIĆ, L. D.; LESJAK, M. M.; ČETOJEVIĆ-SIMIM, D. D. & BEARA, I. N. Activity Profiling of Serbian and Some other European Merlot Wines in Inflammation and Oxidation Processes. **Food Research International**, v. 121, p. 151-160, 2019.

MARÇO, P. H.; VALDERRAMA, P.; ALEXANDRINO, L. G; POPPI, R. J. & TAULER, R. Resolução Multivariada de Curvas com Mínimos Quadrados Alternantes: Descrição, Funcionamento e Aplicações. **Química Nova**, v. 37, n. 9, p. 1525-1532, 2014.

MARTENS, H. & NAES, T. **Multivariate Calibration**, John Wiley & Sons, Chichester, West Sussex, p. 116, 1989.

MATEUS, N.; SILVA, A. M.; RIVAS-GONZALO, J. C.; SANTOS-BUELGA, C. & FREITAS, V. A New Class of Blue Anthocyanin-derived Pigments Isolated from Red Wines. **Journal of Agricultural and Food Chemistry**, v. 51, p. 1919-1923, 2003.

MATEUS, N.; SILVA, A. M. S.; SANTOS-BULEGA, C.; RIVAS-GONZALO, J. C. & de FREITAS, V. Identification of Anthocyanin-flavanol Pigments in Red Wines by NMR and Mass spectrometry. **Journal of Agricultural and Food Chemistry**, v. 50, p. 2110-2116, 2002.

MAXCHEIX, J. J.; FLEURIET, A. & BILLOT, J. **The Main Phenolics of Fruits. In Fruits Phenolics**; CRC Press: Boca Raton, Florida, p. 1-98, 1990.

MELANIE, M. **An Introduction to Genetic Algorithms**. MIT Press, 1998.

MELLO, L. M. R. EMBRAPA, Empresa Brasileira de Pesquisa e Agropécuaria. Uva e Vinho. Informe Técnico. **Panorama da Produção do Vinho no Brasil**, 2017. <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/159111/1/Mello-CampoNegocio-V22-N142-P54-56-2017.pdf>>. Disponível: 09 de janeiro de 2019.

MELLO, L. M. R. **Vitivinicultura Brasileira: Panorama 2012**. Bento Gonçalves: EMBRAPA, Comunicado Técnico, n. 137, p.5, 2013.

MICHALEWICZ, Z. **Genetic Algorithms + Data Structures = Evolution Programs**. (3<sup>rd</sup> Ed.). London, UK, UK: Springer-Verlag, 1996.

MILANEZ, K. D. T. M.; NÓBREGA, T. C. A.; NASCIMENTO, D. S.; INSAUSTI, M.; BAND, B. S. F. & PONTES, M. J. C. Multivariate Modeling for Detecting Adulteration of Extra Virgin Olive Oil with Soybean Oil Using Fluorescence and UV–Vis Spectroscopies: A Preliminary Approach, **LWT - Food Science and Technology**, v. 85, Part A, p. 9-15, 2017.

MILFORD, D. & SANDELL, M. Singular value decomposition using an array of CORDIC processors. **Signal Process**, v. 102, p. 163-170, 2014.

MULLHEN, R. J.; MONEKOSSO, D.; BARMAN, S & REMAGNINO, P. A Review of Ant Algorithms. **Expert Systems With Applications**, v. 36, n. 6, p. 9608-9617, 2009.

MURPHY, K. R.; STEDMON, C. A.; GRAEBER, D. & BRO, R.; Fluorescence Spectroscopy and Multi-way Techniques. PARAFAC. **Analytical Methods**, v. 5, p. 6557-6566, 2013.

MRKUS, L., BATINIĆ, J., BJELIŠ, N. & JAKAS, A. Synthesis and Biological Evaluation of Quercetin and Resveratrol Peptidyl Derivatives as Potential Anticancer and Antioxidant Agents. **Amino Acids**, v. 51, n. 2, p. 319-329, 2019.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

NACHE, M.; SCHEIER, R.; SCHMIDT, H. & HITZMANN, B. Non-invasive Lactate- and pH-monitoring in Porcine Meat Using Raman Spectroscopy and Chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 142, p. 197-205, 2015.

NERY, S. W. L. **Análise de Operadores de Cruzamento Genético Aplicados ao Problema do Caixeiro Viajante**. Monografia. Universidade Federal de Goiás. Catalão, Goiás. 2017.

NETO, J. L. P. S. **Uma Abordagem para Resolver Deficiência de Posto, Desalinhando de Pico e Interferências na Análise de Animas Biogênicas em Peixes Usando HPC-DAD**. Tese de Doutorado. Universidade Federal da Paraíba. João Pessoa, 2018.

NIEDERLE, P.A. **Controvérsias Sobre a Noção de Indicações Geográficas Enquanto Instrumento de Desenvolvimento Territorial: a Experiência do Vale dos Vinhedos em Questão**. In: Congresso Brasileiro de Economia, Administração e Sociologia Rural, 47, Porto Alegre: SOBER, 2009.

NIKOLAJSEN, R. P. H.; BOOKSH, K. S.; HANSEN, A. M. & BRO, R. Quantifying Catecholamines Using Multi-way Kinetic Modelling. **Analytica Chimica Acta**, v. 475, p. 137-150, 2003.

OIV. International Organisation of Vine and Wine. **World Vitiviniculture situation: Statistical Report on World Vitiviniculture**, p. 1-20, 2017.

OLIVIERI, A. & ESCANDAR, G. **Practical Three-way Calibration**, 1<sup>o</sup> ed. Rosario: Elsevier. p. 330, 2014.

OLIVIERI, A. C. Analytical Figures of Merit: From Univariate to Multiway Calibration. **Chemical Review**, v. 114, p. 5358-5378, 2014.

OLIVIERI, A. C.; ARANCIBIA, J. A.; MUNOZ de La PENA, A.; DURÁN-MERÁS, I. & ESPINOSA MANSILLA, A. Second-Order Advantage Achieved with Four-Way Fluorescence Excitation–Emission–Kinetic Data Processed by Parallel Factor Analysis and Trilinear Least-Squares. Determination of Methotrexate and Leucovorin in Human Urine. **Analytical Chemistry**, v. 76, n. 19, p. 5657-5666, 2004.

OLIVIERI, A. C.; WU, H. L. & YU, R. Q. MVC2: a MATLAB Graphical Interface Toolbox for Second-order Multivariate Calibration. **Chemometrics and Intelligent Laboratory Systems**, v. 96, n. 2, p. 246-251, 2009.

PARASTAR, H. & TAULER, R. Multivariate Curve Resolution of Hyphenated and Multidimensional Chromatographic Measurements: A New Insight to Address Current Chromatographic Challenges. **Analytical Chemistry**, v. 86, n. 286, p. 286–297, 2014.

PESSOA, C. M. **Aperfeiçoamento do Algoritmo Colônia de Formigas para o Desenvolvimento de Modelos Quimiométricos**. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul (UFRGS). Porto alegre. p. 131. 2015.

PEÑA, A. M. de la; GOICOECHEA, H. C.; ESCANDAR, G. M. & OLIVIERI, A. C. **Fundamentals and Analytical Applications of Multiway Calibration**. Amsterdam: Elsevier. 2015.

PIRES, D. E. **Uma Proposta de Interdisciplinaridade Utilizando Análise Combinatória e o Algoritmo de Colônia de Formigas no Ensino Médio**. Dissertação de Mestrado. Universidade Federal de Goiás Campos Avançado. Jataí. 2019.

PIEROZAN, V. L. **O Território da Viticultura Orgânica no Território da Vitivinicultura da “Serra Gaúcha”: O Caso dos Viticultores de Cotiporã -RS**. Tese de Doutorado. Universidade Federal do Rio Grande do Sul. Porto alegre, 2017.

PONTES AS, ARAÚJO A, MARINHO W, DINIZ, P. H. G. D.; GOMES, A. A.; GOICOECHEA, H. C.; SILVA, C. E. & ARAÚJO, M. C. U. Ant Colony Optimization for Variable Selection in Discriminant Linear Analysis. **Journal of Chemometrics**, v. 34, n. 12, p. e3292, 2020.

PRATES, A. **Vinho no Egito**. 21 de agosto de 2018. <<https://entrementes.com.br/2018/08/vinho-no-egito/>>. Disponível em: 07 de janeiro de 2019.

RANZAN, C.; STROHM, A.; RANZAN, L.; TRIERWEILER, L. F.; HITZMANN, B & TRIERWEILER, J. O. Wheat Flour Characterization Using NIR and Spectral Filter Based on Ant Colony Optimization. **Chemometrics and Intelligent Laboratory Systems**, v. 32, p.133-140, 2014.

RANZAN, L. **Estudo da Viabilidade do Uso de Espectroscopia Fluorescente 2D para Quantificar Teor de Enxofre em Óleo Diesel**. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul (UFRGS). Engenharia de Sistema. P. 94. 2014.

REAL, M. C. **Os Bons Vinhos do Sul**. (5. Ed). Porto Alegre: Editora Sulina, 1981.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

RIBÉREAU-GAYON, P.; GLORIES, Y.; MAUJEAN, A.; DUBORDIEU, D. **Handbook of enology: The chemistry of Wine Stabilization and Treatments**. 2nd ed., France: John Wiley e Sons. v. 2, p. 441, 2006.

RIVAS, G. J. C.; BRAVO-HARO, S. & SANTOS-BUELGA, C. Detection of compounds Formed Through the Reaction of Malvidin 3-monoglucoside and Catechin in the Presence of Acetaldehyde. **Journal of Agricultural and Food Chemistry**, v. 43, p. 1444–1449, 1995.

RÍOS-REINA, R.; ELCOROARISTIZABAL, S.; OCAÑA-GONZÁLEZ, J. A.; GARCÍA-GONZÁLEZ, D. L.; AAMIGO, J. M. & CALLEJÓN, R. M. Characterization and Authentication of Spanish PDO Wine Vinegars Using Multidimensional Fluorescence and Chemometrics. **Food Chemistry**, v. 230, p. 108-116, 2017.

RODRÍGUEZ-DÍAS, R. C.; AGUILLAR-CABALLOS, M. P. & GÓMEZ-HENS, A. Determination of Some Hydroxybenzoic Acids and Catechins in White Wine Samples by Liquid Chromatography with Luminescence Detection. **Journal of Separation Science**, v. 29, p. 2772-2779, 2006.

RODRÍGUEZ, D. A.; MERÁS, I. D.; DÍAZ, G. T. & WOLD, J. P. Front-face Fluorescence Spectroscopy: A New Tool for Control in the Wine Industry. *Journal of Food Composition and Analysis*, v. 24, p. 257-264, 2011.

RUNGE, J.; HERINGER, O. A., RIBEIRO, J. S. & BIAZATI, L. B. Multi-Element Rice Grains Analysis by ICP OES and Classification by Processing Types. **Food Chemistry**, v. 271, p.419-424, 2019.

RUTAN, S. C.; JUAN, A. & TAULER, R. **Introduction to Multivariate Curve Resolution, in Comprehensive Chemometrics**, Elsevier, Oxford, v. 2, p. 249–259, 2009.

SAAD, R.; BOUVERESSE, D. J.-R.; LOCQUET, N. & RUTLEDGE, D. N. Using pH Variations to Improve the Discrimination of Wines by 3D Front Face Fluorescence Spectroscopy Associated to Independent Components Analysis. **Talanta**, v. 153, p. 278-284, 2016.

SAAVEDRA, J.; FUENTEALBA, C.; YÁÑEZ, L.; BRAVO, M.; Quiroz, W.; LUKACSY, G. & CAROT, J. M. Chemometric Approaches for the Zoning of Pinot Noir Wines from the Casablanca valley, Chile. **Food Chemistry**, v. 127, n. 4, p. 1842-1847, 2011.



SANTOS, D. I.V & GIL, E. S. Fluorimetria na Análise Farmacêutica: uma Revisão. **Revista Eletrônica de Farmácia**, v. 7, n. 1, p. 24-38, 2010.

SANTOS, A. M. P.; LIMA, J. S.; dos SANTOS, I. F.; SILVA, E. F. R.; de SANTANA, F. A.; de ARAUJO, D. G. G. R. & dos SANTOS, L. O. Mineral and Centesimal Composition Evaluation of Conventional and Organic Cultivars Sweet Potato (*Ipomoea batatas* (L.) Lam) Using Chemometric Tools. **Food Chemistry**, v. 273, p. 166-171, 2017.

SARMENTO, M. B. Diagnóstico da Vitivinicultura na Campanha Gaúcha: Análise Swot. **Revista Agropampa**, v. 1, n. 1, p. 65-85, 2016.

SAURINA, J. Characterization of Wines Using Compositional Profiles and Chemometrics. **Trends in Analytical Chemistry**, v. 29, n. 3, p. 234-245, 2010.

SENNA, M. M.; TREVISAN, M. G. & POPPI, R. J. PARAFAC: uma Ferramenta Quimiométrica para Tratamento de Dados Multidimensionais Aplicações na Determinação Direta de Fármacos em Plasma Humano por Espectrofluorimetria. **Química Nova**, v. 28, n. 5, p. 910-920, 2005.

ŠELIH, V. S.; ŠALA, M. & DRGAN, V. Multi-element Analysis of Wines by ICP-MS and ICP-OES and their Classification According to Geographical Origin in Slovenia. **Food Chemistry**, v. 153, p. 414-423, 2014.

SILVA, E. B. **Aplicação do Algoritmo Colônia de Formigas na Restauração de Redes Elétricas de Distribuição**. Dissertação de Mestrado. Universidade Federal da Bahia. Salvador. 2016.

SILVA, S. F. d. **Seleção de Características por Meio de Algoritmos Genéticos para Aprimoramento de Rankings e de Modelos de Classificação**. Tese de Doutorado. Universidade de São Paulo (USP), São Carlos. 2011.

SILVA, M. P. F.; RODRIGUES, L. R. B.; HONORATO, F. A.; PAIM, A. P. S.; PASQUINI, C. & PIMENTEL, M. F. Classification of Gasoline as with or Without Dispersant and Detergent Additives Using Infrared Spectroscopy and Multivariate Classification, **Fuel**, v. 116, p. 151-157, 2014.

SOCHA, K. & DORIGO, M. Ant Colony Optimization for Continuous Domains. **European Journal of Operational Research**, v. 185, n. 3, p. 1155-1173, 2008.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

SOPER, S. A.; MCGOWN, L. B. & WARNER, I. M. Molecular Fluorescence, Phosphorescence, and Chemiluminescence Spectrometry. **Analytical Chemistry**, v. 66, p. 428R– 444R., 1994.

SOUZA, G. G.; MENEGHIN, L. O.; COELHO, S. P.; MAIA, J. F.& SILVA, A. G. A Uva Roxa, *Vitis vinifera* L. (*Vitaceae*) – seus Sucos e Vinhos na Prevenção de Doenças Cardiovasculares. **Natureza On Line**, v. 4, n. 2, p. 80-86, 2006.

SOUZA, F. G. **Método Meta-heurística de Colônia de Formigas e sua Aplicação na Alocação de Petróleo**. Dissertação de Mestrado. Universidade de São Paulo. (USP). São Paulo. 2019.

SUYKENS, J. A. K. & VANDEWALLE, J. (EDS.), **Nonlinear Modeling**, Advanced Black-Box Techniques, Kluwer Academic Publishers, Boston, p. 55-85, 1998.

SHAN, H.; CUIRONG, S. & YUANJIANG, P. Red Wine Polyphenols for Cancer Prevention. **International Journal of Molecular Sciences**, v. 9, p. 842-853, 2008.

SHEN, F.; WU, J.; YING, Y.; LI, B. & JIANG, T. Differentiation of Chinese Rice Wines from Different Wineries Based on Mineral Elemental Fingerprinting. **Food Chemistry**, v. 141, p. 4026-4030, 2013.

SHUKLA, A.; GOUD, V. V. & DAS, C. Antioxidant Potential and Nutritional Compositions of Selected Ginger Varieties Found in Northeast India. **Industrial Crops and Products**, v. 128, p.167-176, 2019.

SKOOG, D. A.; HOLLER, F. J. & CROUCH, S. R. **Princípios de Análise Instrumental**. 6ª ed. Porto Alegre: Editora Bookman, 2009.

SKOOG, D. A.; HOLLER, F. J. & NIEMAN, T. A. **Princípios de Análise Instrumental**, Porto Alegre: Bookman Editora, 2002.

TAULER, R. & BARCELÒ, D. Multivariate Curve Resolution Applied to Liquid Chromatography-diode Array Detection. **Trends Analytical Chemistry**, v. 12, n. 8, p. 319-327, 1993.

TREDOUX, A.; VILLIERS, A.; MÁJEK, P.; LYNEN, F.; CROUCH, A. & SANDRA, P. Stir Bar Sorptive Extraction Combined with GC-MS Analysis and Chemometric Methods for the Classification of South African Wines According to the Volatile Composition. **Journal of Agricultural and Food Chemistry**, v. 56, p. 4286-4296, 2008.

VACLAVIK, L.; LACINA, O.; HAJŠLOVA, J. & ZWEIGENBAUM. The Use of High-performance Liquid chromatography–quadrupole Time-of-flight Mass Spectrometry Coupled to Advanced Data Mining and Chemometric Tools for Discrimination and Classification of Red Wines According to Their Variety. **Analytica Chimica Acta**, v. 685, p. 45-51, 2011.

VAUDOUR, E. The Quality of Grapes and Wine in Relation to Geography: Notions of Terroir at Various Scales. **Journal of Wine Research**, v. 13, p. 117-141, 2002.

VILLIERS, A.; VANHOENACKER, P.; MAJEK, P. & SANDRA, P. Determination of Anthocyanins in Wine by Direct Injection Liquid Chromatography–diode Array Detection–mass Spectrometry and Classification of Wines Using Discriminant Analysis. **Journal of Chromatography A**, v. 1054, p. 195-204, 2004.

VOGEL, A. R.; MENDHAM, J.; DENNEY, R. C.; BARNES, J. D. & THOMAS, M. **Análise Química Quantitativa**, 6ª Ed., Editora LTC, p.462, 2002.

WALKOWIAK, A.; LEDZINSKI, Ł.; ZAPADKA, M. & KUPCEWICZ, B. Detection of Adulterants in Dietary Supplements with *Ginkgo biloba* Extract by Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy and Multivariate Methods PLS-DA and PCA. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 208, p. 222-228, 2019.

WINDIG, W.; GALLAGHER, N. B.; SHAVER, J. M. & WISE, B. M. **Chemometrics and Intelligent Laboratory Systems**, v, 77, n. 1-2, p. 85-96, 2005.

WORKMAN, J.; LAVINE, B.; CHRISMAN, R. & KOCH, M. Process Analytical Chemistry. **Analytical Chemistry**, v. 83, p. 4557-4578, 2011.

WOLD, S.; GELADI, P.; ESBENSEN, K.; ÖHMAN, J. Multi-way principal components- and PLS-analysis. **Journal of Chemometrics**, v. 1, n. 1, p. 41-56, 1987.

WU, H.; LI, Y.; YU, R.; HAI-LONG, W.; LI, Y & RUN-QUIN, Y. Recent Developments of Chemical Multiway Calibration Methodologies with Second-Order or Higher-Order Advantages. **Journal of Chemometrics**, v. 28, p. 476-489, 2014.

YANG, X. S. **Nature-Inspired Optimization Algorithms**. Oxford: Elsevier, 2014.

## REFERÊNCIAS BIBLIOGRÁFICAS

---

YANG, M.; CHEN, Q.; KUTSANEDZIE, F.Y. H., YANG, X.; GUO, Z. & OUYANG, Q. Portable Spectroscopy System Determination of Acid Value in Peanut Oil Based on Variables Selection Algorithms, **Measurement**, v. 103, p. 179-185, 2017.

YASMIN, R. O. **Quantificação de Componentes em Formas Farmacêuticas Sólidas Via Espectroscopia NIR e Resolução Multivariada de Curvas**. Dissertação de Mestrado. Universidade de Lisboa. Lisboa: Portugal. p. 31, 2017.

---

**6 CAPÍTULO - ARTIGO CIENTÍFICO – Southern Brazilian Red Wine  
Geographical Origin by Means of EEM-pH Four Way Data Modelling  
Coupled With One Class Approach**

---

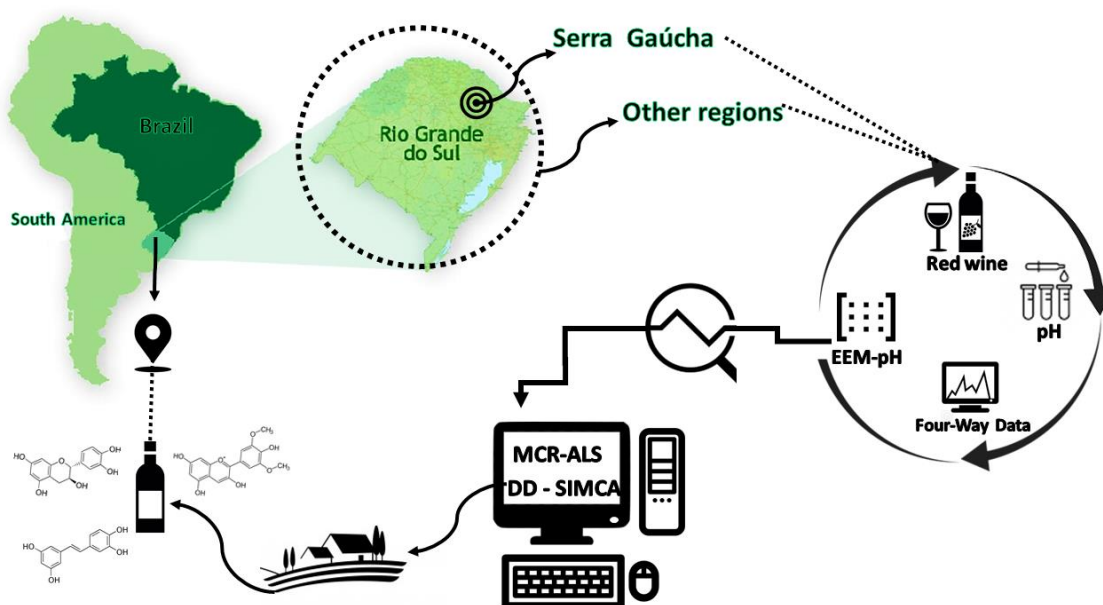


## 6.1 Apresentação

A indicação geográfica do vinho está normalmente relacionada à sua qualidade e ao seu valor comercial. No Rio Grande do Sul, existem duas áreas que possuem um grande potencial na produção de vinhos tintos finos.

Para auxiliar na rastreabilidade da indicação geográfica dos vinhos tintos produzidos no estado do Rio Grande do Sul, com distintas variedades de uvas, foram selecionadas as regiões da Serra Gaúcha e da Campanha.

Uma metodologia analítica por espectroscopia de fluorescência foi desenvolvida empregando o uso de quimiometria para avaliar os compostos químicos presentes no vinho tinto produzidos nessas regiões.



## 6.2 Artigo

Este artigo, foi aceito para o periódico *Food Chemistry* e trata do desenvolvimento de um método analítico para rastreabilidade de vinhos tintos finos brasileiros por Espectroscopia de Fluorescência molecular e análise multivariada de dados de ordem superior.





## Southern Brazilian Red Wine Geographical Origin by Means of EEM-pH Four Way Data Modelling Coupled With One Class Approach

*Layane Lenardon Vinciguerra<sup>a</sup>, Fernanda Carla Böck<sup>b</sup>, Mateus Pires Schneider<sup>b</sup>, Natalia Alejandra Pisoni Canedo Reis<sup>a</sup>, Leticia Flores Silva<sup>c</sup>, Kelly Christina Mendes de Souza<sup>d</sup>, Celito Crivellaro Guerra<sup>c</sup>, Adriano de Araújo Gomes<sup>b,d\*</sup>, Ana Maria Bergold<sup>a</sup> and Marco Flôres Ferrão<sup>b,e\*\*</sup>*

<sup>a</sup>Faculty of Pharmacy, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil.

<sup>b</sup>Institute of Chemistry, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil.

<sup>c</sup>Empresa Brasileira de Pesquisa Agropecuária, Centro Nacional de Pesquisa de Uva e Vinho, Bento Gonçalves, Rio Grande do Sul, Brazil.

<sup>d</sup>Instituto de Ciências Exatas-ICE. Universidade Federal do Sul e Sudeste do Pará- Unifesspa.

<sup>e</sup>Instituto Nacional de Ciência e Tecnologia-Bioanalítica (INCT-Bioanalítica), Cidade Universitária, Zeferino Vaz s/n, Campinas, São Paulo, Brazil.

### ABSTRACT

Excitation Emission data recorded at deferments pH has been exploited by means Multivariate curve resolution with alternative least squares in order to find out qualitative information about Brazilian red wines. In addition, the geographical traceability of wines produced in the state of Rio Grande do Sul from the most important production area Serra Gaúcha, was studied by means of one class approach, namely DD-SIMCA. In this study, 53 samples from target class (Serra Gaúcha) and 20 from other producing regions. The fluorescence signal recorded corresponds to nine emission (51 variables) excitation (12 variables) matrices (EEM) recorded at different pH (3 up 11) generating four-way data. By MCR-ALS

decomposition, eight factors were retrieved and related to typical chemical compounds found in red wine. In addition, the EEM-pH data have used to build classification one class model taking into account MCR scores, and good correct rating rate has been achieved.

**Keywords:** *Multiway Data; MCR-ALS; Red wine; One Class Classification; Serra Gaúcha.*

## 1. INTRODUCTION

Red wine is one of the oldest beverages in the world, being appreciated and cultivated today (JOHSON, 1989), it has stood out more and more in the consumer market, due to its different aromas, textures and flavors from the grape varieties and the winemaking process. For this reason, it is also one of the most famous drinks in the world, which can have its price ranging from very cheap to thousands of dollars per bottle (MARISA, ALMEIDA & VASCONCELOS, 2003; ÁLVAREZ *et al.*, 2007). All wine characteristics depend on the environment in which the grapes are grown (DULEY, 2021).

Geographical origin usually adds to the commercial value of a wine; some regions are famous for producing excellent wines with high commercial value. The origin of a product is an important issue for the food industry, not only for consumers but also for producers and distributors (LUYKX & VAN RUTH, 2008). In Brazil, the state of Rio Grande do Sul stands out as a wine-producing region and the Serra Gaúcha (SG) as the most important grape producing areas including fine red wine (*Vitis vinifera*) (ALMEIDA *et al.*, 2016; GIULIANI, 2007). The Serra Gaúcha is a mountainous region and is cut by the 29<sup>st</sup> parallel with altitudes from 500 to 800 m. It has a precipitation index of about 1.800 mm/year and has a humid climate with acidic and sandy soils, rich in organic matter that has little efficient drainage. Despite the problems of excess rain during the months of grape harvest season, the climate and methods of planting used in Serra Gaúcha make it possible to produce grapes with regular high acidity (FARIAS, 2014; SOUZA *et al.*, 2006; GUERRA *et al.*, 2005; FLORES & MEDEIROS, 2013).

Due to its popularity, wine has become the target of investigation because of fraud including adulteration, false age declaration and false indication of geographical origin. Much effort has been invested to the development of methods for assessing wine quality, to avoid fraud and to indicate its true geographical origin (ÁLVAREZ *et al.*, 2007). The literature has a number of reports of analytical approaches to establish the composition and authenticity of the wine according to its geographical origin and grape variety. For instance, among more conventional approaches there are liquid and gas chromatography coupled to

several detectors (mass (MS), diode array (DAD), MS-MS and fluorescence) (BALLABIO, SKOV, LEARDI & BRO, 2008; VILLIERS, VANHOENACKER, MAJEK & SANDRA, 2004; DIMITROVSKA, TOMOVSKA & BOCEVSKA, 2013; RODRÍGUEZ, AGUILAR-CABALLOS & GÓMEZ-HENZ, 2006) and atomic spectrometry, such as inductively coupled plasma mass spectrometry (ICP-MS), inductively coupled plasma optical emission spectrometry (ICP-OES) (SHEN *et al.*, 2013; ŠELIH, SALA & DRGAN, 2014). These analytical techniques are effective in elucidating the organic and elemental composition of wines but require steps of sample preparation, costly equipment and high maintenance, which may represent a limiting factor for small industries, producers and consumers.

In addition, alternative methods have been carried out using spectroscopic techniques coupled to chemometric tools. These include proton nuclear magnetic resonance ( $^1\text{H}$  RMN) (ANASTASIADI *et al.*, 2009; LEE *et al.*, 2009), near infrared spectroscopy (NIR) (COZZOLINO, SMYTH, & GISHEN, 2003; LIU *et al.*, 2008), Fourier transform infrared with attenuated total reflectance spectroscopy (FTIR-ATR) (SCHMIDTKE, SMITH, MÜLLER, & HOLZAPFEL, 2012), capillary electrophoresis (CE-DAD) (RODRÍGUEZ, DURÁN-MERÁS, GALEANO-DÍAZ & WOLD, 2011), UV-Vis spectroscopy (AZCARATE *et al.*, 2015) and fluorescence spectroscopy (RODRÍGUEZ, DURÁN-MERÁS, GALEANO-DÍAZ & WOLD, 2011; AZCARATE *et al.*, 2015; SAAD, BOUVERESSE, LOCQUET & RUTLEDGE, 2016). Mainly due to the resulting lack of a need for a sample preparation step, the combination of spectroscopic techniques and chemometrics have advantages such as relative higher speed, and lower cost of execution, are non-destructive and/or non-invasive, have low reagent consumption and, at last but not least can be used on site (GREDILLA *et al.*, 2016).

In the specific case of fluorescence spectroscopy for wine analysis, some advantages of this method, like high sensitivity and selectivity, deserve attention when compared with other spectroscopic approaches (LAKOWICZ, 2006). Fluorescent properties of a specific chemical compound may change as a function of the pH. For example, anthocyanins and polyphenols present in wine affect the

recording of the fluorescence signal at different pH values, allowing a better characterization of the fluorescent compounds present in wines (SAAD, BOUVERESSE, LOCQUET & RUTLEDGE, 2016; GREDILLA *et al.*, 2016; LAKOWICZ, 2006).

Moreover, the versatility of the instruments lies on in its the different forms of signal recording: a simple fluorescence value obtained for a specific excitation and emission wavelength; emission spectra with a sample excited at a fixed excitation wavelength; and a fluorescence map, corresponding to several emission spectra obtained by distancing excitation wavelengths (LAKOWICZ, 2006; KUMAR & MISHRA, 2012). These types of data are arranged in an excitation-emission matrix (EEM).

The present his work is devoted to the development of a fluorescent pH-dependent fingerprint of the Brazilian red wines, based on 3D front face fluorescence and extended Multivariate curve resolution with alternative least squares (e-MCR-ALS) (AZCARATE, GOMES, LA PEÑA & GOICOECHEA, 2018), for the classification proposes with respect to geographical origin considering the most famous region as target class, the Serra Gaúcha, by means of one class method Data driven soft independent modelling of class analogy (DD-SIMCA). This approach is an important contribution to characterization and quality control of red wines. Furthermore, the work addresses an important yet little exploited chemometric field, for-way data modeling in classification problems.

## 2. EXPERIMENTAL

### 2.1. Reagents and materials

Buffer solutions were prepared from pH values 3 to 11, by using the following reagents: potassium dihydrogen phosphate ( $\text{KH}_2\text{PO}_4$ ) from Êxodo Científico, Hortolândia, SP, Brazil; Phosphoric acid ( $\text{H}_3\text{PO}_4$ ) from Nuclear, Diadema, SP, Brazil; Acetic acid ( $\text{CH}_3\text{COOH}$ ) from Química Moderna, Barueri, SP, Brazil; and sodium acetate ( $\text{CH}_3\text{COONa}\cdot 3\text{H}_2\text{O}$ ), anhydrous bibasic potassium phosphate ( $\text{K}_2\text{HPO}_4$ ), sodium bicarbonate ( $\text{NaHCO}_3$ ), and sodium carbonate

(Na<sub>2</sub>CO<sub>3</sub>) from Dinâmica, Indaiatuba, SP, Brazil. Ultrapure water (Resistivity > 18MΩ MilliQ, Millipore) was used for solution preparations.

## 2.2. Wine samples

Seventy-three samples were used, of which 53 came from Serra Gaúcha and 20 from the others regions also from same state. The samples included ten red grape varieties: Cabernet Sauvignon (CS), Pinot Noir (PN), Merlot (MT), Marselan (MN), Teroldego (TO), Tannat (TT), Malbec (MC), Cabernet Franc (CF), Ancellotta (AA) and Petit Syrah (PS) from harvested in 2017 and 2018. For each pH (3-11), buffer solutions were adjusted using a bench pHmeter (Ultra Denver Instrument<sup>®</sup>). 450 μL of sample were added to 15 mL of buffer, for each pH. All samples were donated by *EMBRAPA Uva e Vinho / Bento de Gonçalves (Rio Grande do Sul)* and kept at room temperature.

## 2.3. EEM measurement

Fluorescence spectral measurements were performed using a Spectrophotometer Varian Cary Eclipse, with two Czerny-Turner monochromators and a xenon flash lamp connected to a PC microcomputer via a serial RS232 (GPIB) interface. The excitation and emission monochromator slit widths were both set to 10 nm. A 600 V PMT Detector with Cary Eclipse software was used to acquire the data. The EEM measurements were recorded as a set of excitation spectra from 244 to 340 nm and emission wavelengths of 368 to 468nm.

## 2.4. Chemometric procedures

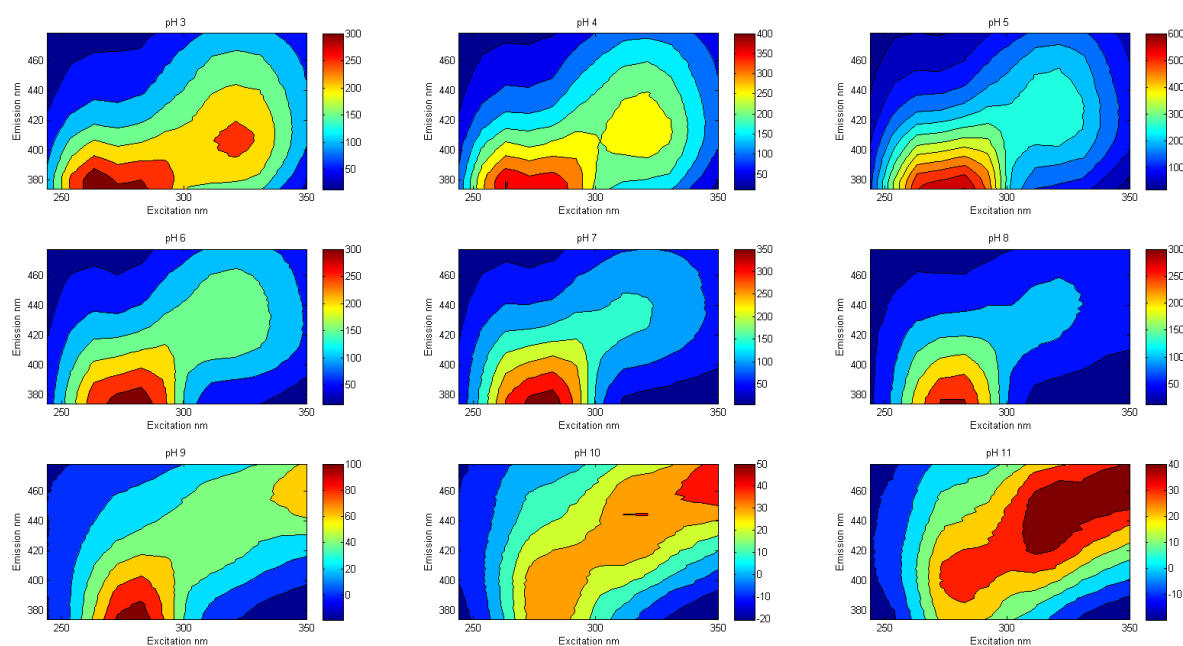
The fluorescent excitation emission matrix (EEM) records at different pH were arranged in augmented matrix (**Daug**) in column wise direction. For each sample, 9 EEM were obtained (pH 3 up 11). These EEM, for the same sample, were arranged in an unfolded matrix (Excitation × Emission-pH) sized 12 × 447. All 73 arrays were stacked forming **Daug** (876 × 447). **Figure 1S** shows a schematic of how the input matrix for the MCR was generated. All chemometric calculations were carried out in a Matlab<sup>®</sup> version 7.1.0 (The MathWorks Inc.,

MA, USA) environment. MCR-ALS and pattern recognition models were obtained by means of MCR ALS toolbox (JAUMOT & TAULER, 2010) (<http://www.mcrals.info/>) and DD-SIMCA freely available at Github (<https://github.com/yzontov/dd-simca.git>) (ZONTOV, RODIONOVA, KUCHERYAVSKIY & POMERANTSEV, 2017) respectively. For classification models, the pool of samples was partitioned through the Kennard-Stone method using a homemade Matlab routine.

### 3.0. Results and Discussion

#### 3.1. General comments

Red wine, in addition to water and ethanol, contains a complex mixture of organic compounds resulting in an average pH of about 3.5. Under this chemical condition, several compounds are present in the protonated form. However, increasing the pH to near neutral values or even strongly basic promotes deprotonation, thus modifying the fluorescence profile of the wine constituents. The anthocyanins, responsible for main wine color, are in equilibrium, for example the flavinic cation ( $AH^+$ )/quinoidal base (A), carbinol (B)/chalcone (C) (SAAD, BOUVERESSE, LOCQUET & RUTLEDGE, 2016; TÔRRES *et al.*, 2011). Changes in pH favors some of the species by increasing or decreasing the fluorescence signal (see **Figure 1**).



**Figure 1:** Typical fluorescence landscape at different pH values.

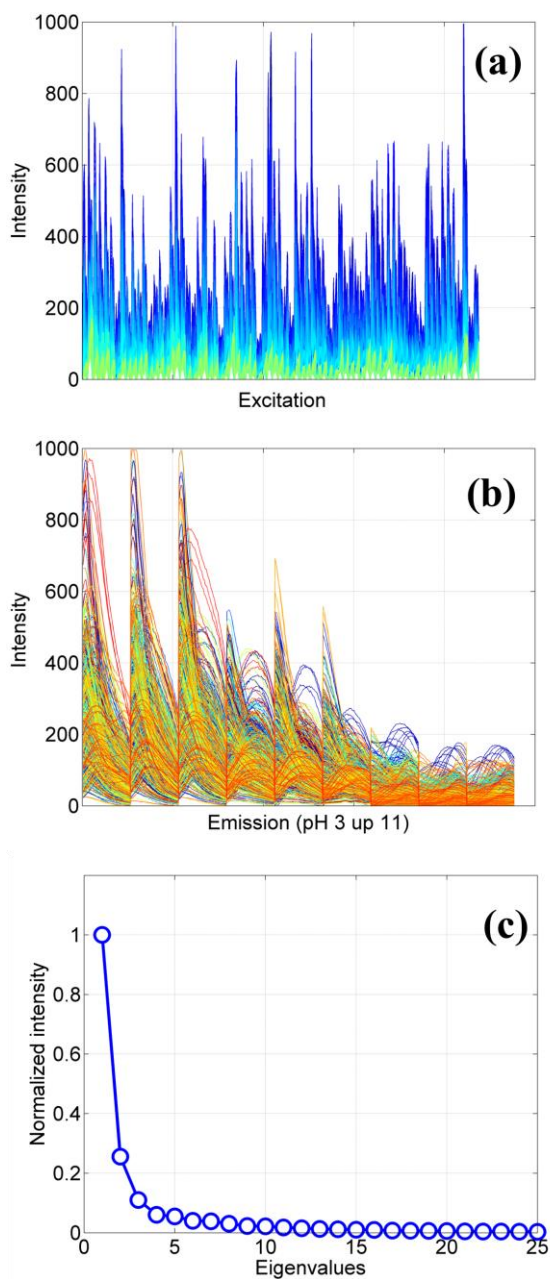
As can be seen, by comparing the fluorescence maps obtained at different pH values, but still in an acid medium, there is a predominance of protonated species like cation ( $AH^+$ ) and carbinol (B). These compounds exhibit the maximum emission observed for smaller wavelengths (see **Figure 1**). The pH detachments for higher values increase chalcone and quinone contents characterized by blue–violet color that the wine starts to exhibit (see **Figure 2S**). In addition, the maximum values of fluorescence intensity occur at longer wavelengths, denoting less energetic transitions.

In a strongly basic medium, a brown color is predominant, due to the binding of anthocyanins to other phenolic compounds, a process called co-pigmentation (SAAD, BOUVERESSE, LOCQUET & RUTLEDGE, 2016; TÔRRES *et al.*, 2011; BOULTON, 2001; MURIAS *et al.*, 2005; GUERRERO *et al.*, 2010). These compounds show very different fluorescence when compared with substances that predominate in acid medium. The sensitivity to pH of the chemical constituents of the wine, with respect to its fluorescent properties, can then be exploited with purpose of distinction based on geographical origin and grape variety.



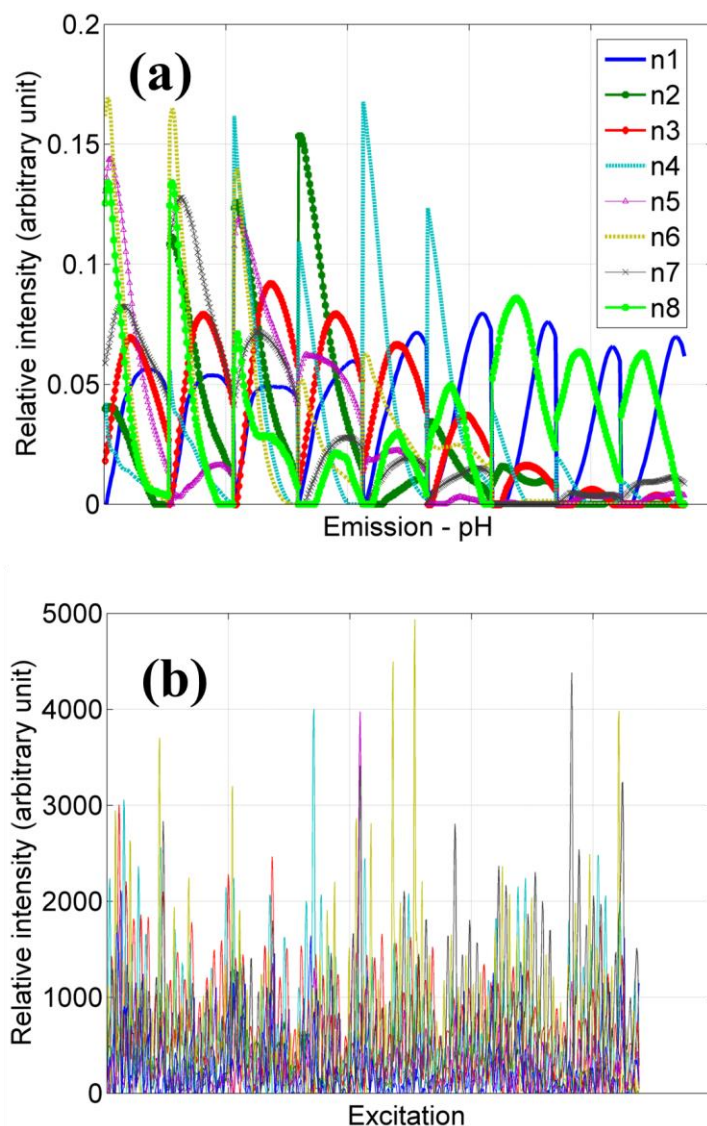
### 3.2. MCR decomposition and feasibility solution

In **Figure 2a** and **2b** is depicted **Daug** matrix in column and row space respectively. Recorded excitation profiles for the 73 samples are displayed In **Figure 2a** and in **Figure 2b** shows the emission profiles for each investigated pH value. Here it is also possible to visualize the changes in the formation and intensity of the spectra due to the increase in pH.



**Figure 2:** Augmented data plot: (a) augmented mode (excitation) and (b) unfolded mode pH  $\times$  emission and in (c) eigenvalue of **Daug** data.

Estimation of the MCR factors was carried out by means of **Daug** matrix singular value decomposition (SVD) and that approach indicated eight factors, as can you be seen in **Figure 2C**. Initial estimate was obtained for the mode emission-pH (non-augmented mode) by means pure variable detection. Alternative least squares optimization was carried out under non-negativity constrain in both directions. It reached convergence after 51 iterations. The final solution corresponds to 98.9% of explained variance and lack of fit (LOF) equal to 10.4%. This value is very small, considering PCA LOF (9.3%). **Figure 3** depict de MCR optimized profiles.



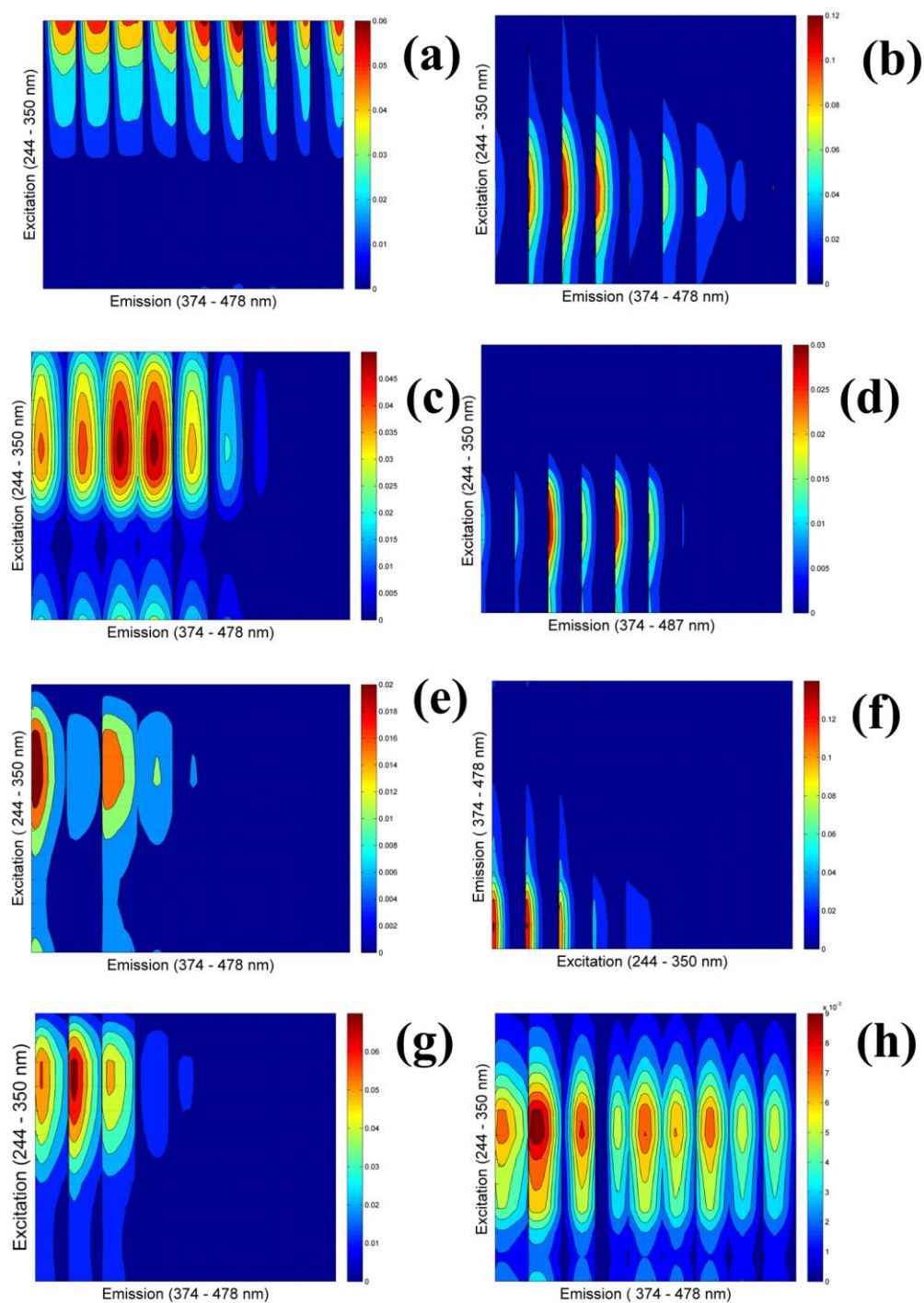
**Figure 3:** Retrieve profiles by MCR-ALS in (a) unfolded mode and (b) augmented mode.

As can you be seen in **Figure 3a**, both shape and intensity of retrieved emission profiles change when pH increase. Some profiles are more intensity in acid pH, in other hand other ones in basic pH. This behavior allows finding more information about wine samples. In **Figure 3b** is shown excitation retrieved profiles that also change but inter samples and its intensity (or area) is directly proportional to concentration.

In order to investigate the feasibility of MCR solution the magnitude of the rotational ambiguity that affects received profiles have been examined by calculation of the relative maximum ( $f_{max}$ ) and minimum ( $f_{min}$ ) contribution of every component in the total signal (GUERRERO et al., 2010). The  $f_{max}$  and  $f_{min}$  values are displayed in **Table 1S**.

Note that the differences between  $f_{max}$  and  $f_{min}$  are very close to zero, this suggests that the solution found is practically free of rotational ambiguity. This is due to the use of fluorescence matrices, for a same sample, having been recorded at different pH values drastically reducing the possibility of rotating the MCR solution without losing fit.

In addition, the profiles retrieved by the MCR provide valuable information about the samples. The **Figure 4** shows how each of the eight profiles changes in shape and intensity with increasing pH.



**Figure 4:** Fluorescence surfaces recovered for each factor by MCR-ALS (letters a, b, c, d, e, f, g and h correspond to factors 1, 2, 3, 4, 5, 6, 7 and 8 respectively).

In general, samples show lower fluorescence intensities at high pH values (neutral or basic medium for example). In an attempt to find a correspondence

between profiles recovered by MCR-ALS and chemical compounds present in wine, the results presented in **Table 1** were found.

**Table 1:** Attempt to assign profiles retrieved via MCR-ALS based on previous literature data.

MCR Retrived Profile	Ex (nm)	Em (nm)	Assignment	Fluorescent Properties	Reference
1	350	432	Fluorescente Back ground	-	-
2	280	<368	Flavonoids like Catechin and Epicatechin	(278-360)	Rodríguez-Delgado <i>et al.</i> , 2001 Airado-Rodríguez <i>et al.</i> , 2009 Rodríguez <i>et al.</i> , 2011
3	253 / 321	403	Phenolic acids like vanillic acids	241, 282, 305 354	Rodríguez-Delgado <i>et al.</i> , 2001 Airado-Rodríguez <i>et al.</i> , 2009 Rodríguez <i>et al.</i> , 2011
4	272	<368	Phenolic acids like gallic and	(318 -385) (300/330-370/392)	Rodríguez-Delgado <i>et al.</i> , 2001 Airado-Rodríguez <i>et al.</i> , 2009 Rodríguez <i>et al.</i> , 2011
5	321	384	Phenolic acids liek caffeic acid, ferulic acid and p-coumaric acid	262 -426 260-400 260-422	Rodríguez-Delgado <i>et al.</i> , 2001 Airado-Rodríguez <i>et al.</i> , 2009 Cabrera-Bañegil <i>et al.</i> , 2017 Rodríguez <i>et al.</i> , 2011
6	263	380	Anthocyanins like Cyanidin and Malvidin Glycosides	280 - 355	Rodríguez-Delgado <i>et al.</i> , 2001 Airado-Rodríguez <i>et al.</i> , 2009 Rodríguez <i>et al.</i> , 2011 Forino <i>et al.</i> , 2019
7	263 / 321	394	Stilbenoid glucoside like <i>t</i> -Resveratrol and <i>t</i> -Piceid	(290–300/390–395) (318 -385) (300/330-370/392)	Rodríguez-Delgado <i>et al.</i> , 2001 Höfener <i>et al.</i> , 2013 Airado-Rodríguez <i>et al.</i> , 2009 Rodríguez <i>et al.</i> , 2011
8	301	388	Unknown	(278-360)	-

The first one was attributed to the fluorescence background; note that it is practically constant for all pH values (see **Figure 4a**). The second (**Figure 4b**) was attributed to fluorescence of flavonoids like catechin and epicatechin, both compounds exhibit very similar excitation and emission profiles, including Cabrera-Bañegil, Hurtado-Sánchez, Galeano-Díaz, & Durán-Merás, (2017), reported the joint quantification of these two polyphenols by fluorescence and PARAFAC. This corroborates the idea of a single MCR factor associated with these two compounds. Phenolic acids like vanillic, gallic and caffeic/ferulic/ p-coumaric were related to profiles three (**Figure 4c**), four (**Figure 4d**) and five (**Figure 4e**) respectively. Factor six (**Figure 4f**), in turn, corresponds to the typical fluorescence signal from anthocyanins like cyanidin and malvidin glycosides, both abundant in red wines. The seventh (**Figure 4h**) profile recovered by the MCR is related to stilbenoid glucoside like trans-resveratrol and trans-piceid. Nothing related to profile eight (**Figure 4h**) was found in the literature.

### *3.3 Geographical traceability*

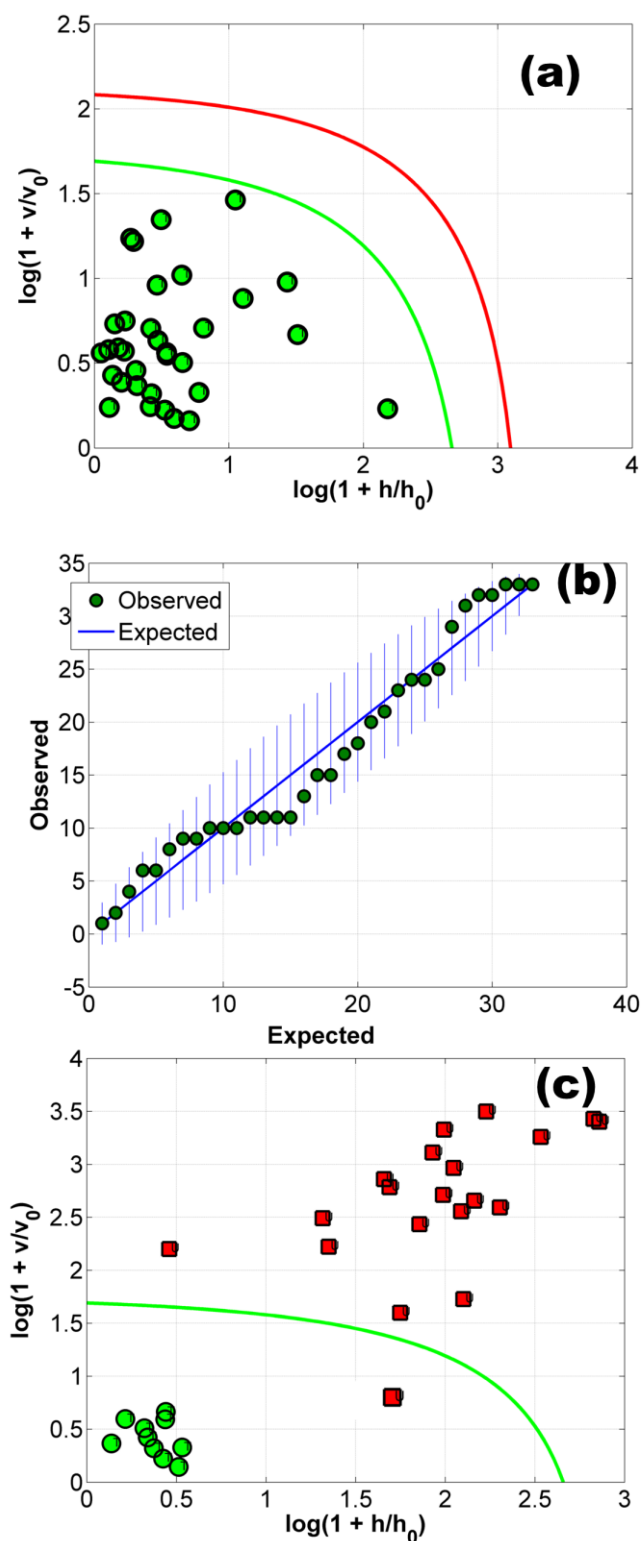
In order to build method for recognizing the authenticity of samples that belongs to Serra Gaúcha a one class approach based on DD-SIMCA was fitted on MCR score. The set of samples of the target class were partitioned in trainee (33 samples), validation (10 samples) and test (10 samples) via Kennard-Stone algorithm. To the test group were added the 20 samples not belonging to the target class. Different models were fitted considering different numbers of principal components and applied to predict samples from the validation set. And the results are shown in **Table 2**.

**Table 2:** Statistical summary of fit and prediction.

	PC	Regular	Extreme	Outlier	<sup>a</sup> Sensitivity
Training set (33) <sup>b</sup>	2	32	1	0	0.97
	3	33	0	0	1
	4	32	1	0	0.97
	5	32	1	0	0.97
	PC	Regular	Extreme	Outlier	<sup>a</sup> Sensitivity
Validation set (10) <sup>c</sup>	2	10	0	0	1
	3	10	0	0	1
	4	10	0	0	1
	5	10	0	0	1

<sup>a</sup>Sensitivity= 100%(Samples – Extremes)/Samples. <sup>b</sup>Traning Samples and <sup>c</sup>Validation Samples.

As can be seen in **Table 2**, for all numbers of PCs evaluated, maximum sensitivity was achieved in the validation set. However only for 3 CP there was no extreme sample in the training set. Therefore, 3 PC was selected for the final model. And in **Figure 5a** and **Figure 5b** are shown the acceptance area for the training set and the extreme plot respectively.



**Figure 5:** DD-SIMCA results: (a) acceptance area established for training set (green line is the border of regular samples and red line is the border of outlier samples. Sample between both lines is called extreme samples. In (b) Extreme plot and (c) results of applying the model to the test set (green circles are samples of the target class and red squares are samples not belonging to the target class). Specificity= 100% \*(External objects)/ Samples.



As can be seen in **Figure 5a**, all training samples showed status of rules being positioned within the boundary delimited by the established green line considering  $\alpha$  of 0.01. In addition, the absence of an outlier was confirmed by the extreme plot (see **Figure 5b**). Note that all samples are contained in the elliptical region.

When this model was applied to the test set containing both samples of the target class and also not belonging to the target class, their position with respect to the decision boundary can be seen in **Figure 5c**. All samples of the target class were properly recognized as belonging to the target class with maximal sensitivity equal to 1, samples of the non-target class were also adequately rejected by the model and the specificity found was 0.97.

#### **4.0 CONCLUSIONS**

This work demonstrated the modeling of four-way data generated by the recording of EEM at different pH by MCR-ALS in order to recover the contributions of the main fluorescent compounds present in the red wine. And good agreement between the results and the literature data were found. The scores of the MCR model were used in the built of a model for authenticating samples from the main wine producing region of Brazil, Serra Gaúcha, based on a one class approach called DD-SIMCA. The model exhibits a high capacity to recognize samples of the target class as well as to reject samples not belonging to Serra Gaúcha. Therefore, this methodology, in addition to extracting valuable quantitative information, was useful in the task of geographic authentication.

#### **ACKNOWLEDGEMENTS**

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. The authors are thankful to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for financial support, scholarships and research grants. The authors, also,

are thankful to Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) Uva e Vinho from Bento Gonçalves and Associação Brasileira de Enologia (ABE) for donations of wine samples.

The English text of this paper has been revised by Sidney Pratt, Canadian, MAT (The Johns Hopkins University), RSA dip - TESL (Cambridge University).

## REFERENCES

- AIRADO-RODRÍGUEZ, D., GALEANO-DÍAZ, T., DURÁN-MERÁS, I. & WOLD, J. P. (2009). Usefulness of Fluorescence Excitation-Emission Matrices in Combination with PARAFAC, as Fingerprints of Red Wines. *Journal of Agricultural and Food Chemistry*, 57, 1711-1720. <https://doi.org/10.1021/jf8033623>.
- ALMEIDA, C. M. S., ALMEIDA, A. C., GODOY, M. L. D. P., PIERRE, T. D. S. & GODOY, J. M. (2016). Differentiation among Brazilian wine regions based on lead isotopic data. *Journal of the Brazilian Chemical Society*, 27(6), 1026-1031. <https://doi.org/10.5935/0103-5053.20150358>.
- ÁLVAREZ, M., MORENO, I. M., JOS, A. M., CAMEÁN, A. M. & GONZÁLEZ, A. G. (2007). Differentiation of ‘two Andalusian DO ‘fino’ wines according to their metal content from ICP-OES by using supervised pattern recognition methods. *Microchemical Journal*, 87, 72-76. <https://doi.org/10.1016/j.microc.2007.05.007>.
- ANASTASIADI, M., ZIRA, A., MAGIATIS, P., HAROUTOUNIAN, S. A., SKALTSOUNIS, A. L. & MIKROS, E. (2009). <sup>1</sup>H NMR-Based metabolomics for the classification of greek wines according to variety, region, and vintage. Comparison with HPLC data. *Journal of Agricultural and Food Chemistry*, 57(23), 11067-11074. <https://doi.org/10.1021/jf902137e>.
- AZCARATE, S. M., GOMES, A. A., ALCARAZ, M. R., DE ARAÚJO, M. C. U., CAMIÑA, J. M. & GOICOECHEA, H. C. (2015). Modeling excitation–

- emission fluorescence matrices with pattern recognition algorithms for classification of Argentine white wines according grape variety. *Food Chemistry*, 184, 214-219. <https://doi.org/10.1016/j.foodchem.2015.03.081>.
- AZCARATE, S., GOMES, A. A., DE LA PEÑA, A. M. & GOICOECHEA, H. C. (2018). Modeling second-order data for classification issues: data characteristics, algorithms, processing procedures and applications. *Trends in Analytical Chemistry*, 107, 151-168. <https://doi.org/10.1016/j.trac.2018.07.022>.
- BALLABIO, D., SKOV, T., LEARDI, R. & BRO, R. (2008). Classification of GC-MS measurements of wines by combining data dimension reduction and variable selection techniques. *Journal of Chemometrics*, 22, 457-463. <https://doi.org/10.1002/cem.1173>.
- BOULTON, R. (2001). The copigmentation of anthocyanins and its role in the color of red wine: A critical review. *American Journal of Enology and Viticulture*, 52(2), 67-87.
- CABRERA-BAÑEGIL, M., HURTADO-SÁNCHEZ, M. C., GALEANO-DÍAZ T. & DURÁN-MERÁS, I. (2017). Front-face fluorescence spectroscopy combined with second-order multivariate algorithms for the quantification of polyphenols in red wine samples. *Food Chemistry*, 220, 168-176. <http://doi.org/10.1016/j.foodchem.2016.09.152>.
- COZZOLINO, D., SMYTH, H. E. & GISHEN, M. (2003). Feasibility study on the use of visible and near-Infrared spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. *Journal of Agricultural and Food Chemistry*, 51(26), 7703–7708. <https://doi.org/10.1021/jf034959s>.
- DIMITROVSKA, M., TOMOVSKA, E. & BOCEVSKA, M. (2013). Characterisation of Vranec, Cabernet Sauvignon and Merlot wines based on their chromatic and anthocyanin profiles. *Journal of the Serbian Chemical Society*, 78(9), 1309-1322. <https://doi.org/10.2298/JSC130101026D>.

- DULEY, G., DUJOURDY, L., KLEIN, S., WERWEIN, A., SPARTZ, C., GOUGEON, R. D. & TAYLOR, D. K. (2021). Regionality in Australian Pinot noir wines: A study on the use of NMR and ICP-MS on commercial wines. *Food Chemistry*, 340, <https://doi.org/10.1016/j.foodchem.2020.127906>.
- FARIAS, R. (2014). *Curso para profissionais ciclo I*. Rio de Janeiro: Associação Brasileira de Sommeliers.
- FORINO, M., GAMBUTI, A., LUCIANO, P. & MOIO, L. (2019). Malvidin-3-O-glucoside Chemical Behavior in the Wine pH Range. *Journal of Agricultural and Food Chemistry*, 67, 1222-1229. <https://doi.org/10.1021/acs.jafc.8b05895>.
- FLORES, S. S. & MEDEIROS, R. M. V. (2013). Ruralidades na compreensão dos territórios do vinho e sua identidade. *Revista de Geografia Agrária*, 8(15) 1-19.
- GIULIANI, E. (2007). The selective nature of knowledge networks in clusters: evidence from the wine industry. *Journal of Economic Geography*, 7, 139-168. <https://doi.org/10.1093/jeg/lbl014>.
- GUERRA, C. C., MANDELLI, F., TONIETTO, J., ZANUS, M. C. & CAMARGO, U. A. (2005). *Conhecendo o essencial sobre uvas e vinhos*. Bento Gonçalves: Embrapa Uva e Vinho.
- GUERRERO, R. F., PUERTAS, B., JIMÉNEZ, M. J., CACHO, J. & CANTOS-VILLAR, E. (2010). Monitoring the process to obtain red wine enriched in resveratrol and piceatannol without quality loss. *Food Chemistry*, 122, 195-202. <https://doi.org/10.1016/j.foodchem.2010.02.057>.
- GREDILLA, A., DE VALLEJUELO, S. F. O., ELEJOSTE, N., DE DIEGO, A. & MADARIAGA, J. M. (2016). Non-destructive Spectroscopy combined with chemometrics as a tool for Green Chemical Analysis of environmental samples: A review. *Trends in Analytical Chemistry*, 76, 30-39. <https://doi.org/10.1016/j.trac.2015.11.011>.

- HÖFENER, S., KOOIJMAN, P. C., GROEN, J., ARIESEB, F. & VISSCHER, L. (2013). Fluorescence behavior of (selected) flavonols: a combined experimental and computational study. *Physical Chemistry Chemical Physics*, *15*, 12572-12581. <https://doi.org/10.1039/c3cp44267e>.
- JAUMOT, J. & TAULER, R. (2010). MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution. *Chemometrics Intelligent Laboratory Systems*, *103*, 96-107. <http://doi.org/0.1016/j.chemolab.2010.05.020>.
- JOHSON, H. (1989). *The Story of Wine*. London: Mitchell-Beazley.
- KUMAR, K. & MISHRA, A. K. (2012). Application of parallel factor analysis to total synchronous fluorescence spectrum of dilute multifluorophoric solutions: Addressing the issue of lack of trilinearity in total synchronous fluorescence data set. *Analytica Chimica Acta*, *755*, 37-45. <https://doi.org/10.1016/j.aca.2012.10.024>.
- LAKOWICZ, J. R. (2006). *Principles of fluorescence spectroscopy* (3th ed.). Baltimore, Maryland: USA: Springer, 954.
- LEE, J. E., HWANG, G. S., BERG, F. V. D., LEE, C. H. & HONG, Y. S. (2009). Evidence of vintage effects on grape wines using <sup>1</sup>H NMR-based metabolomic study. *Analytica Chimica Acta*, *648*, 71-76. <https://doi.org/10.1016/j.aca.2009.06.039>.
- LIU, L., COZZOLINO, D., CYNKAR, W. U., DAMBERGS, R. G., JANIK, L., O'NEILL, B. K., COLBY, C. B. & GISHEN, M. (2008). Preliminary study on the application of visible–near infrared spectroscopy and chemometrics to classify Riesling wines from different countries. *Food Chemistry*, *106*, 781-786. <https://doi.org/10.1016/j.foodchem.2007.06.015>.
- LUYKX, D. M. A. M. & VAN RUTH, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, *107*(2), 897-911. <https://doi.org/10.1016/j.foodchem.2007.09.038>.

- MARISA, C., ALMEIDA, R. & VASCONCELOS, M. T. S. D. (2003). Multielement composition of wines and their precursors including provenance soil and their potentialities as fingerprints of wine origin. *Journal of Agricultural and Food Chemistry*, 51(16), 4788-4798. <https://doi.org/10.1021/jf034145b>.
- MURIAS, M., JÄGER, W., HANDLER, N., ERKER, T., HORVATH, Z., SZEKERES, T., NOHL, H. & GILLE, L. (2005). Antioxidant prooxidant and cytotoxic activity of hydroxylated resveratrol analogues: structures-activity relationship. *Biochemical Pharmacology*, 69(6), 903-912. <https://doi.org/10.1016/j.bcp.2004.12.001>.
- RODRÍGUEZ, R. C., AGUILAR-CABALLOS, M. P. & GÓMEZ-HENZ, A. (2006). Determination of some hydroxybenzoic acids and catechins in white wine samples by liquid chromatography with luminescence detection, *Journal of Separation Science*, 29, 2772-2779. <https://doi.org/10.1002/jssc.200600206>.
- RODRÍGUEZ, D. A., DURÁN-MERÁS, I., GALEANO-DÍAZ, T. & WOLD, J. P. (2011). Front-face fluorescence spectroscopy: A new tool for control in the wine industry. *Journal of Food Composition and Analysis*, 24, 257-264. <https://doi.org/10.1016/j.jfca.2010.10.005>.
- RODRÍGUEZ-DELGADO, M. A, MALOVANA, S., PÉREZ, J. P., BORGES, T. & GARCÍA MONTELONGO, F. J. (2001). Separation of phenolic compounds by high-performance liquid chromatography with absorbance and fluorimetric detection. *Journal of Chromatography A*, 912, 249-257. [https://doi.org/10.1016/s0021-9673\(01\)00598-2](https://doi.org/10.1016/s0021-9673(01)00598-2).
- SAAD, R., BOUVERESSE, D. J. R., LOCQUET, N. & RUTLEDGE, D. N. (2016). Using pH variations to improve the discrimination of wines by 3D front face fluorescence spectroscopy associated to independent components, *Talanta*, 153, 278-284. <https://doi.org/10.1016/j.talanta.2016.03.023>.

- ŠELIH, V. S., SALA, M. & DRGAN, V. (2014). Multi-element analysis of wines by ICP-MS and ICP-OES and their classification according to geographical origin in Slovenia. *Food Chemistry*, 153, 414-423. <http://doi.org/10.1016/j.foodchem.2013.12.081>.
- SOUZA, G. G., MENEGHIN, L. O., COELHO, S. P., MAIA, J. F. & SILVA, A. G. (2006). A uva roxa, *Vitis vinífera L.* (Vitaceae) – seus sucos e vinhos na prevenção de doenças cardiovasculares. *Natureza On Line*, 4(2), 80-86.
- SCHMIDTKE, L. M., SMITH, J. P., MÜLLER, M. C. & HOLZAPFEL, B. P. (2012). Rapid monitoring of grapevine reserves using ATR-FT-IR and chemometrics. *Analytica Chimica Acta*, 732, 16-25. <https://doi.org/10.1016/j.aca.2011.10.055>.
- SHEN, F., WU, J. YIBIN, Y., LI, B. & JIANG, T. (2013). Differentiation of Chinese rice wines from different wineries based on mineral elemental fingerprinting. *Food Chemistry*, 141, 4026-4030. <http://doi.org/10.1016/j.foodchem.2013.06.119>.
- TÔRRES, A. R., LYRA, W. S., ANDRADE, S. I. E., ANDRADE, R. A. N., DA SILVA, E. C., ARAÚJO, M. C. U. & GAIÃO, E. N. (2011). A digital image-based method for determining of total acidity in red wines using acid-base titration without indicator. *Talanta*, 84(3) 601-606. <https://doi.org/10.1016/j.talanta.2011.02.002>.
- VILLIERS, DE A., VANHOENACKER, G., MAJEK, P. & SANDRA, P. (2004). Determination of anthocyanins in wine by direct injection liquid chromatography–diode array detection–mass spectrometry and classification of wines using discriminant analysis. *Journal of Chromatography A*, 1054, 195-204. <https://doi.org/10.1016/j.chroma.2004.07.087>.
- ZONTOV, Y.V., RODIONOVA, O.Y., KUCHERYAVSKIY, S.V. & POMERANTSEV, A.L. (2017). DD-SIMCA – A MATLAB GUI tool for data driven SIMCA approach. *Chemometrics Intelligent Laboratory Systems*, 167, 23-28. <http://doi.org/10.1016/j.chemolab.2017.05.010>.









O pH do vinho tinto é em média de cerca de 3,5, devido a sua mistura complexa de compostos orgânicos. Nessa condição química, vários compostos estão presentes na forma protonada. No entanto, o aumento do pH para valores quase neutros ou mesmo fortemente básicos promove a desprotonação, modificando assim o perfil de fluorescência dos constituintes do vinho. A mudança de pH para valores mais altos aumenta os teores de chalcona e quinona, caracterizados pela cor azul-violeta que o vinho começa a exibir (**Figura 2S** – do anexo (material suplementar do artigo)).

As antocianinas, responsáveis pela principal cor do vinho, encontram-se em equilíbrio. Por exemplo o cátion flavínico (AH<sup>+</sup>) / base quinóide (A), carbinol (B) / chalcona (C) (SAAD, BOUVERESSE, LOCQUET & RUTLEDGE, 2016; TÔRRES *et al.*, 2011). Mudanças no pH favorecem algumas espécies, aumentando ou diminuindo o sinal de fluorescência. A sensibilidade ao pH dos constituintes químicos do vinho, no que diz respeito às suas propriedades fluorescentes, pode então ser explorada com finalidade de distinção com base na origem geográfica e no varietal empregado.

Dessa forma, diferentes sinais de fluorescência foram analisados através dos perfis de excitação/emissão registrados para as 73 amostras, para cada valor de pH investigado. Aqui também foi possível visualizar as mudanças na formação e intensidade dos espectros, devido ao aumento do pH (**Figura 1** – do artigo).

Na tentativa de encontrar uma correspondência entre os 8 perfis recuperados pelo MCR-ALS e os compostos químicos presentes no vinho, foram comparados os perfis de fluorescência recuperados com dados da literatura (**Tabela 1** – do artigo).

O primeiro foi atribuído ao fundo de fluorescência; observe que é praticamente constante para todos os valores de pH. O segundo foi atribuído à fluorescência de flavonóides como catequina e epicatequina, ambos os compostos exibem excitação e perfis de emissão muito semelhantes, segundo Cabrera-Bañegil, Hurtado-Sánchez, Galeano-Díaz, & Durán-Merás, (2017), que relatam a quantificação conjunta desses dois polifenóis por fluorescência e PARAFAC. Isso confirma a ideia de um único fator de MCR associado a esses dois compostos. Ácidos fenólicos como vanílico, gálico e caféico / ferúlico / p-cumárico foram relacionados aos perfis três (**Figura 4c** – do artigo), quatro (**Figura 4d** – do artigo) e cinco, respectivamente. O fator seis, por sua vez, corresponde ao sinal de fluorescência típico de antocianinas como glicosídeos de cianidina e malvidina, ambos abundantes em vinhos tintos. O sétimo perfil recuperado pelo MCR

está relacionado ao glicosídeo estilbenoide como o trans-resveratrol e o trans-piceide. Porém, nada relacionado ao perfil oito foi encontrado na literatura.

Estudos recentes, tem comparado o desempenho dos algoritmos ACO e GA na seleção de variáveis, tanto para o desenvolvimento de modelos multivariados de regressão, quanto para a classificação. Como exemplo de regressão empregando mínimos quadrados parciais (PLS), podemos citar o desenvolvimento de modelos PLS com dados NIR, para a determinação do índice de acidez do óleo de amendoim, onde ambos os algoritmos (ACO e GA) apresentaram baixos valores para o erro quadrático médio de previsão (RMSEP). Neste trabalho, o melhor modelo PLS-GA selecionou 20 variáveis resultando num RMSEP de 0,3200, enquanto que o PLS-ACO selecionou 101 variáveis resultando num RMSEP de 0,3260, ambos empregando 7 variáveis latentes (YANG *et al.*, 2017).

Já o trabalho apresentado por Pontes e colaboradores (2020), compara o desempenho do ACO e do GA visando o desenvolvimento de modelos de análise discriminante. Dois estudos de caso foram apresentados: (i) classificação de óleos vegetais comestíveis (em relação ao óleo de base) por espectrometria ultravioleta-visível (UV-Vis) e (ii) classificação simultânea de amostras de chá com relação para tipo e origem geográfica via espectrometria de infravermelho próximo (NIR). Para os modelos de classificação dos óleos vegetais comestíveis, a seleção de variáveis através dos algoritmos ACO e GA resultaram em taxas de acerto de 100 e 95,7 %, respectivamente. Por outro lado, para os modelos de classificação das amostras de chás, a seleção de variáveis através dos algoritmos ACO e GA resultaram em taxas de acerto de 92 e 84 %, respectivamente (PONTES *et al.*, 2020).

Por outro lado, alguns trabalhos publicados apresentam comparativos do desempenho do SW com outras técnicas de seleção de variáveis, buscando o desenvolvimento de modelos de classificação por análise linear discriminante. Um estudo realizado para avaliar o envelhecimento da cerveja, empregou GA e SW para selecionar informações relevantes (variáveis) a partir de espectros no NIR, além da ortogonalização supervisionada de Gram-Schmidt. Avaliando apenas os resultados por validação cruzada, os pesquisadores obtiveram 100% de acerto tanto para o LDA-GA quanto para o LDA-SW, porém os algoritmos diferiram nas variáveis selecionadas para a construção dos modelos LDA (GHASEMI-VARNAMKHAZI & FORINA, 2014).

Em outro trabalho, modelos LDA foram desenvolvidos empregando para seleção de variáveis os algoritmos GA, SW e SPA (algoritmo das projeções sucessivas), utilizando espectros de infravermelho (MIR e NIR) do resíduo de destilação para classificação de amostras de gasolina em dois grupos: com ou sem aditivos (detergentes ou dispersantes). Neste caso, os melhores resultados foram obtidos usando LDA-GA ou SPA-LDA para a região do infravermelho médio (SILVA *et al.*, 2014).

Todavia, poucos trabalhos apresentam a aplicação dos algoritmos de seleção de variáveis para dados de fluorescência. No trabalho proposto por Milanez e colaboradores, é apresentado um estudo comparativo de métodos quimiométricos usados para quantificar a adulteração do azeite de oliva extravirgem (EVOO) com óleo de soja comestível por fluorescência e espectroscopias de UV-Vis. Diferentes estratégias de regressão multivariada foram avaliadas: mínimos quadrados parciais (PLS) usando espectro total; PLS com coeficientes de regressão significativos selecionados pelo algoritmo Jack-Knife (PLS-JK) e regressão linear múltipla (MLR) com seleção prévia de variáveis por algoritmos *stepwise* (MLR-SW); algoritmo de projeções sucessivas (MLR-SPA); e algoritmo genético (MLR-GA). Foi avaliada a capacidade preditiva dos modelos, para cada técnica espectroscópica. Para a espectroscopia de fluorescência, foram obtidos resultados de predição satisfatórios para todos os modelos de regressão com valores de RMSEP variando de 14,0 a 17,5 g/kg, sendo que o modelo MLR-SW apresentou o menor erro. Quando os métodos de regressão foram avaliados para espectros de UV-Vis, valores mais elevados de RMSEP foram encontrados, variando de 13,3 a 30,4 g/kg. Além disso, no teste F, houve diferença entre os métodos PLS, PLS-JK e os métodos MLR desenvolvidos com seleção prévia das variáveis, para o qual os modelos PLS apresentaram menor erro (MILANEZ *et al.*, 2017).

Ao compararmos os desempenhos dos algoritmos de seleção de variáveis descritos na literatura para dados espectroscópicos, podemos observar que não existe o algoritmo perfeito e que sempre apresente os melhores resultados. O desempenho destes pode depender da técnica espectroscópica empregada, da natureza das amostras, ou até mesmo do critério de classificação que se deseja (procedência geográfica, tipo, adulteração, presença de aditivos, entre outros).



---

## **8 CONCLUSÃO**

---





Neste trabalho foi possível desenvolver metodologias analíticas para identificação da origem geográfica de vinhos tintos produzidos no Rio Grande do Sul, empregando dados de espectrometria por fluorescência e ferramentas quimiométricas.

Na primeira parte, foram inicialmente comparados os desempenhos dos métodos de seleção de variáveis: algoritmo colônia de formigas (ACO), algoritmo genético (GA) e formulação *stepwise* (SW). Para tanto foram desenvolvidos modelos de análise discriminante linear, empregando as variáveis espectrais dos pH3 e pH7, além da fusão deste, a partir do desdobramento dos dados das matrizes de excitação/emissão visando gerar dados de primeira ordem.

Quando comparamos os algoritmos ACO e GA, todos os modelos utilizando o pH3 e pH7, apresentaram melhor desempenho para as amostras da Serra Gaúcha do que para a Campanha devido ao efeito do desbalanceamento que afetou a construção dos modelos de classificação. Contudo, quando realizado o modelo SW para a fusão dos dados com os dois pHs (pH3+pH7), foi possível compensar o desbalanceamento obtendo resultados satisfatórios para as duas regiões estudadas, tanto para as amostras de treinamento quanto para a teste, resultando num modelo com taxa de acerto superior a 90%. Este modelo, com certeza poderia auxiliar no controle de qualidade para reconhecimento da origem geográfica de amostras de vinho tinto provenientes das regiões da Campanha e Serra Gaúcha.

Na segunda parte do trabalho foi desenvolvido a modelagem de dados por quatro vias que foram geradas através de matrizes de excitação/emissão (EEM) pelo registro obtido por meio da análise de espectrofluorimetria em diferentes pH e por resolução multivariada de curvas por mínimos quadrados alternantes (MCR-ALS). Com isso, foi possível recuperar as contribuições dos principais compostos fluorescentes do vinho tintos brasileiros (Serra Gaúcha e Campanha). Dessa forma, obteve-se uma boa concordância entre os resultados da análise quando comparamos com os dados da literatura.

Com os resultados dos dados tratados no MCR-ALS, foi desenvolvido um modelo para a certificação de origem geográfica, para amostras da principal região vinícola do Brasil, a Serra Gaúcha, através do método *data driven soft independent modelling of class analogy* (DD-SIMCA). O modelo apresenta alta capacidade de reconhecimento de amostras da classe alvo, bem como de rejeição de amostras não pertencentes à Serra

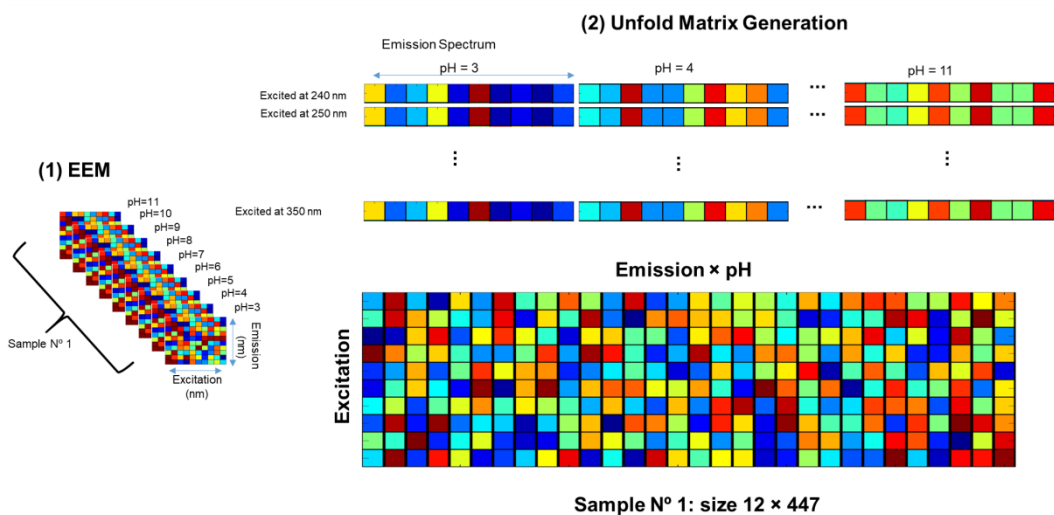
## CONCLUSÃO

---

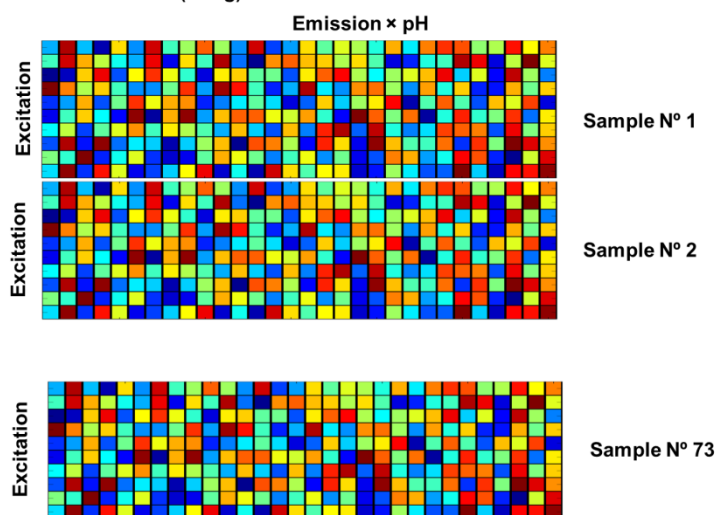
Gaúcha. Portanto, esta metodologia, além de extrair valiosas informações quantitativas, foi útil na tarefa de autenticação geográfica.



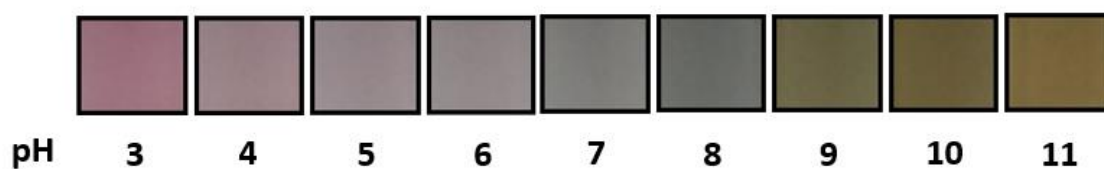




**(4) Augmented Matrix Generation (Daug) /Column wise mode**



**Figure 1S:** Data organization scheme.



**Figure 2S:** Change in wine color due to pH variation.

**Table 1S:** Feasibility MCR solution.

<b>Components</b>	$f_{\min}$	$f_{\max}$
<i>n1</i>	0.1492	0.1492
<i>n2</i>	0.1115	0.1115
<i>n3</i>	0.2488	0.2488
<i>n4</i>	0.2555	0.2555
<i>n5</i>	0.1445	0.1445
<i>n6</i>	0.2676	0.2676
<i>n7</i>	0.2850	0.2850
<i>n8</i>	0.1017	0.1017

---