Authentication of Australian Red Wines Using Fluorescence Spectroscopy and Machine Learning Classification

Ranaweera Kaluarachchige Ruchira Ranaweera

B.Sc.(Hons), M.Sc.

A thesis submitted for the degree of Doctor of Philosophy

School of Agriculture, Food and Wine

The University of Adelaide

Adelaide, Australia

November 2021

To my loving daughters,
Ayana and Kiara

Contents

Thesis Summary i

Declaration v

Publications vi

Conferences viii

Panel of Supervisors ix

Acknowledgements x

Thesis Structure xii

Chapter 1 Authenticity and traceability in the wine industry 01

Statement of Authorship 02

Publication - Invited Book Chapter 04

Chapter 2 A Review of wine authentication using spectroscopic approaches in combination with chemometrics 33

Statement of Authorship 34

Publication - Review Paper 36

Research Questions 51

Summary of Aims 54

Contents

Chapter 3	**Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling**	**56**
	Statement of Authorship	**57**
	Publication	**59**
Chapter 4	**Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine**	**71**
	Statement of Authorship	**72**
	Publication	**74**
Chapter 5	**Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine**	**90**
	Statement of Authorship	**91**
	Publication	**93**
Chapter 6	**Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines**	**129**
	Statement of Authorship	**130**
	Publication	**132**

Contents

Chapter 7 Concluding Remarks & Future Perspectives **149**

7.1. Conclusions **150**

7.1.1. Development of the method using fluorescence spectroscopy for wine authentication **150**

7.1.2. Verification of the method according to region and exploring the varietal authentication with insight into chemical markers **151**

7.1.3. Investigation of intraregional variation of Shiraz wines using A-TEEM and machine learning modelling **152**

7.1.4. Tracing the molecular fingerprint of wine throughout the winemaking process and recognising the blending percentage of variatal wine **153**

7.2. Future perspectives **155**

List of Abbreviations **157**

List of Figures **159**

List of Tables **166**

Apendix A **170**

Apendix B **191**

Thesis Summary

Verification of geographical origin, grape variety, and year of production of wine is essential in validating quality, identifying fraud, and improving the economic value of wine according to those important extrinsic factors. The identity of a wine is influenced mainly by its origin, as reflected in a wine's composition. Therefore, analytical methods that identify authentication markers to discriminate wine according to the origin (or other variables) are required.

Over the years, numerous methods for wine authentication have been identified, from traditional analytical methods to rapid advanced instrumental techniques. However, there is a lack of a robust but simple technique that gives rapid results and is sensitive enough to discriminate wines accurately. This forms the topic of the thesis, which begins with a published book chapter that covers current aspects of wine authenticity and traceability in terms of technological and consumer perspectives **(Chapter 1)**. Different spectroscopic approaches and chemometric methods used in wine authentication in the past decades have been evaluated for their characteristics in the next chapter, published as a review paper **(Chapter 2).**

As a rapid, straightforward, selective, and sensitive method that yields a molecular fingerprint of wine, fluorescence spectroscopy was identified upon reviewing the literature as a promising method to investigate wine authentication. Several original research studies were subsequently performed with the aim of understanding the potential of applying fluorescence spectroscopy in combination with multivariate data analysis for wine authentication **(Chapters 3 to 6).** Finally, the conclusions and future directions of the study are included in the final chapter **(Chapter 7).**

In the initial research publication using spectrofluorometric analysis **(Chapter 3)**, a method based on absorbance-transmission and fluorescence excitation-emission matrix (known as the A-TEEM technique) was

investigated as a tool for regional authentication of commercial Australian Cabernet Sauvignon wines from three different Geographical Indications (GIs) in comparison to wines from Bordeaux, France as an international benchmark. The potential of A-TEEM spectroscopy for wine authentication was assessed in comparison to elemental profiling using inductively coupled plasma-mass spectrometry (ICP-MS) as a reference method for geographical authentication. Among other multivariate algorithms used for classification of the wines, a novel machine learning technique known as extreme gradient boosting discriminant analysis (XGBDA) yielded 100 % correct classification for all tested regions using the fluorescence data, and overall 97.7 % for ICP-MS. This result emphasised the possibility of applying A-TEEM and XGBDA for accurate authentication of wines.

With these encouraging GI authentication results, a further study was undertaken to verify the origin of wine according to both geographical and varietal variations. A wide range of commercially-produced but unreleased wines from ten different Australian GIs and three varieties (Shiraz, Cabernet Sauvignon, and Merlot) were studied in the second research publication **(Chapter 4)**. This study identified the effectiveness of combining absorbance and fluorescence data from A- TEEM as a multi-block data set to maximise the model's robustness. Excellent results were obtained in relation to cross-validation for each class (100 % for variety and 99.7 % for region of origin), again highlighting the effectiveness of A- TEEM data with XGBDA. In addition, A-TEEM data was interrogated using partial least squares regression (PLSR) models to rapidly quantify 24 phenolic compounds of relevance to red wine (i.e., anthocyanins, flavonols, flavan-3-ols, hydroxycinnamates). Principal component analysis of the phenolic compound concentrations revealed differences among the varieties and regions, helping to understand the chemical markers that were important in classification. These findings further strengthen the potential of using the A-TEEM technique for differentiation of wine, not only from GIs at state level but also those from adjacent regions such as Clare, Barossa, and Eden Valleys within a state.

Further testing the A-TEEM technique for its ability to discriminate wine at a sub-regional level, research-scale and commercial unreleased Shiraz wines from five different areas within the Barossa Valley GI along with Eden Valley GI were analysed to explore their intra-regional variations. The samples were from three consecutive years, which allowed for authentication testing according to the vintage, as reported in the third original research study submitted for publication **(Chapter 5).** The sensitivity of the A-TEEM technique allied with XGBDA facilitated 100 % accuracy in classifying Shiraz wines according to the sub-region of origin and year of production. Additionally, A-TEEM data were modelled with PLSR in comparison to reference method data to predict basic chemical parameters of the samples (i.e., pH, alcohol %v/v, titratable acidity), which enhances the utility of the A-TEEM technique as a rapid method for deployment in the wine industry.

In wine authentication, it is important to understand the impact of winemaking processes on chemical markers at different stages of production. Hence, variations in molecular fingerprint of wines throughout the process such as after primary fermentation, after malolactic fermentation, and before blending were determined with the A-TEEM technique. XGBDA discriminated wines according to their origin (variety and region) with 100 % accuracy, eliminating the influence of stage of processing on spectral signature. Also, blending different grape varieties or wine from different GIs (as permitted by relevant regulations) is crucial in winemaking. However, it is important to determine whether blending a small proportion (up to 15 % of other varietal or regional wine as per Wine Australia regulations) can be detected for authentication purposes. Unreleased commercially-produced monovarietal wines were prepared with a series of blends containing Shiraz with Cabernet Sauvignon and Shiraz with Grenache and analysed with regression. XGB regression precisely predicted the percentage in the blend, achieving R^2 CV of 1.00 and RMSECV of 0.00028 in comparison to PLSR, which did not perform as well. The results of this final study of the thesis were submitted for publication as a short communication **(Chapter 6)**.

In summary, this PhD thesis has been devoted to the development of a rapid analytical method to accurately authenticate wine according to geographical origin, variety, and vintage. The use of absorbance and/or fluorescence spectroscopy in conjunction with machine learning classification proved to be highly promising for this purpose. The outcomes of this thesis not only contribute to enriching scientific research but also offer opportunities for potential commercial application in the wine industry as a powerful tool for wine analysis, and in particular, validation of origin and composition.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

The author acknowledges that copyright of published works contained within the thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Ranaweera Kaluarachchige Ruchira Ranaweera Date: 15 November 2021

Publications

This PhD thesis includes a published book chapter (Comprehensive Foodomics, Elsevier), and a collection of three papers that were published in peer-reviewed scientific journals (Food Chemistry, impact factor 7.514 and Molecules, Impact factor 4.411) during candidature (Chapter 2 to 4). Chapter 5 and 6 include manuscripts that have been submitted to Food Chemistry and OENO One, respectively.

Chapter 1 **Ranaweera, R. K. R**, Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Authenticity and traceability in the wine industry: From analytical chemistry to consumer perceptions. In A. Cifuentes (Ed.), Comprehensive Foodomics (pp. 452-480) Oxford: Elsevier.

Chapter 2 **Ranaweera, R. K. R.,** Capone, D. L., Bastian, S. E. P., Cozzolino, D., & Jeffery, D. W. (2021). A review of wine authentication using spectroscopic approaches in combination with chemometrics. Molecules, 26(14), 4334.

Chapter 3 **Ranaweera, R. K. R.,** Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. Food Chemistry, 335, 127592.

Chapter 4 **Ranaweera, R. K. R.,** Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. Food Chemistry, 130149.

Chapter 5 **Ranaweera, R. K. R.,** Bastian, S. E. P., Gilmore, A. M., Capone, D. L., & Jeffery, D. W. (2021). Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine. Food Chemistry, Submitted.

Chapter 6 Ranaweera, R. K. R., Gilmore, A. M., Bastian, S. E. P., Capone, D. L., & Jeffery, D. W. (2021). Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines. OENO One, Submitted.

Additionally, the following publications included in the Appendix section were produced during candidature:

Appendix A Souza Gonzaga, L., Bastian, S. E. P., Capone, D. L., **Ranaweera, R. K. R.**, & Jeffery, D.W. (2021). Modelling Cabernet Sauvignon wine sensory traits with spectrofluorometric data. OENO One, 55 (4), 4805.

Appendix B **Ranaweera, R. K. R.** , Gilmore, A. M., Capone, D. L., Bastian, S. E. P., Jeffery, D. W. (2020). Authenticating the geographical origin of wine using fluorescence spectroscopy and machine learning. Paper presented at the XIIIth International Terroir Congress, Adelaide, Australia.

Conferences

1 **17th Australian Wine Industry Technical Conference (AWITC)**, Adelaide, Australia, 21 - 24 July 2019. Poster presentation: "Authentication of Cabernet Sauvignon wines from different regions of Australia"

2 **School of Agriculture, Food & Wine Annual Postgraduate Symposium,** Adelaide, Australia, 24 – 25 September 2019. Oral presentation: "Chemical markers for authentication of Australian wines"

3 **XIII th International Terroir Congress**, Online platform, 17 - 18 November 2020. Oral presentation: "Application of fluorescence spectroscopy with multivariate analysis for authentication of Shiraz wines from different regions"

4 **Macrowine 2021,** Online platform, 23 - 30 June 2021.
Oral presentation: "Wine authentication with fluorescence spectroscopy and XGBoost"

5 **Crush 2021- The Grape and Wine Science Symposium**, Adelaide, Australia, 16 June2021. Oral Presentation: "Authenticating the origin of wine: Fingerprinting with fluorescence spectroscopy in conjunction with machine learning"

Co-authored conference abstracts:

6 **70th ASEV National Conference,** Napa, California USA, 17– 20 June 2019. Technical Abstracts: "Analyses of smoke-tainted wine using simultaneous absorbance- transmittance and fluorescence excitation-emission mapping"

7 **Macrowine 2021,** Online platform, 23 – 30 June 2021.
Poster presentation: "Chemical and sensory diversity of regional Cabernet Sauvignon wines"

8 **WAC 2022: 5th International Conference on Active Wine Compounds.**
Accepted poster presentation: "Regional classification of Cabernet Sauvignon wines using extreme gradient boost (XGB) analysis of simultaneous absorbance-transmittance fluorescence excitation-emission (A- TEEM) spectroscopy"

Panel of supervisors

Principal Supervisor Associate Professor **David W. Jeffery**

0000-0002-7054-0374
The University of Adelaide
School of Agriculture, Food and Wine

Co-Supervisor Dr. **Dimitra L. Capone**

0000-0003-4424-0746
The University of Adelaide
School of Agriculture, Food and Wine

Co-Supervisor Associate Professor **Susan E. P. Bastian**

0000-0002-8790-2044
The University of Adelaide
School of Agriculture, Food and Wine

Independent Advisor Associate Professor **Daniel Cozzolino**

0000-0001-6247-8817
University of Queensland
Brisbane, QLD, AU

Acknowledgements

Successful completion of this project could not have been accomplished without the involvement, assistance, and contribution of so many people whose names may not all be enumerated.

In the first place, I would like to express my deep gratitude to my principal supervisor, Associate Professor David Jeffery, for giving me the opportunity to carry out the research and providing me with indispensable supervision throughout this endeavor. Your vision, integrity, enthusiasm towards the subject deeply inspired me. Thank you for being the best mentor, understanding my strengths, and guiding me towards success. Thank you for your endless support and care every day from the start of the project to date. I feel so fortunate to learn a heap from you and forever grateful to you.

I would like to acknowledge my sincere gratitude to my co-supervisors, Associate Professor Susan Bastian and Dr. Dimitra Capone, for all the guidance, expert advice and for sharing your knowledge all over the research project to achieve this successful completion. Thank you for being an amazing role model and giving me all the care and encouragement throughout this journey. I also thank Associate Professor Daniel Cozzolino, my independant advisor, for all your support and expert advice on this project and for giving me the opportunity to publish my literature review. I am extremely grateful to Dr. Adam Gilmore, HORIBA Instruments Inc USA, for your enormous inputs and expert advice for this project on multivariate analysis and the Aqualog analysis. Moreover, I would like to thank Warren Birchmore at Accolade Wines, Anthony Quinn and Andrew Jane at Quark Photonics and Chris Henderson at Jeffress Engineering for their generous in-kind contributions to the project.

A special thanks need to give to my parents and my sister for your unconditional love, understanding, and empathy. Whenever and wherever, you are always there for me. Everything I have and everything I am, I owe it all to you. Also, I am extending my heartiest thank you to my husband and two daughters for your continued and unfailing

love, precious support, sacrifice and understanding during the pursuit of my Ph.D. I could have never accomplished this without you.

I want to thank my colleagues for all the support and encouragement throughout my PhD study, in particular, Dr. Chen Liang, Dr. Liang Chen, Dr. Ahn Nguyen, Dr. Carolyn Puglisi, Dr. Renata Restic, Dr. Lukas Danner, Nick Van Holst, Joanne Hards, Lira Souza Gonzaga, Claire Armstrong, Xingchen Wang, Ross Sanders, Judith Unterkofler, Hugh Holds, Yihe Sui, Colleen Szeto and Pietro Previtali at the University of Adelaide. It's a great pleasure to work with all of you. Also, I would like to thank Dr. Erandi Hewawasam for your encouragement and inspiration to start and continue my PhD.

I am immensely appreciative to the School of Agriculture Food and Wine at The University of Adelaide for the financial support provided through the Australian Government Scholarship and to Wine Australia for Wine Australia supplementary scholarship.

Thesis Structure

The thesis is composed of chapters including publications and manuscripts. As outlined below, it begins with a published book chapter consisting of an overview of the research area followed by a review paper, and four original research papers. Finally, conclusions and future directions of the study are contained within the last chapter.

Chapter 1 The first chapter comprises a broad view of literature related to the wine authenticity and traceability within the scope of analytical techniques and consumer perspectives.

Chapter 2 This chapter specifically explores the spectroscopic approaches as rapid techniques that have been applied in wine authentication in the past decades. Their characteristics were reviewed along with the application of chemometrics. At the end of the chapter, research questions and a summary of research aims are included.

Chapter 3 Application of fluorescence spectroscopy in conjunction with chemometrics to build robust authentication models for wine was investigated using A-TEEM technique and compared to the effectiveness of ICP-MS as a reference method.

Chapter 4 A combination of absorbance and EEM data from A-TEEM technique was analysed in a multi-block setting with XGBDA for regional and varietal authentication of Australian red wine. To understand the chemical markers relevant for classification, phenolic concentration predictions were carried out.

Chapter 5 The ability of A-TEEM technique to discriminate wine at a sub-regional level was explored using Shiraz wines from five different areas within the Barossa Valley GI, South Australia. Further authentication according to the vintage was conducted as the wines were produced in three consecutive years.

Chapter 6 As a preliminary study, the impact of winemaking processes on the molecular fingerprint of wine at different stages of production and the blending percentages of different wines were identified by applying A-TEEM technique.

Chapter 7 Overall conclusions were drawn by considering the research studies conducted based on A-TEEM technique in combination with machine learning. Future perspectives of wine authenticity were discussed in terms of the potential of employing the A-TEEM method in the wine industry.

Nothing in life is to be feared, it is only to be understood.

Now is the time to understand more,

so that we may fear less.

- Marie Curie

Chapter 1

Invited Book Chapter

Publication

Authenticity and Traceability in the Wine Industry: From Analytical Chemistry to Consumer Perceptions.

Ranaweera, R.K.R., Souza Gonzaga, L., Capone, D.L., Bastian, S.E.P., & Jeffery, D.W. (2021). Authenticity and traceability in the wine industry: From analytical chemistry to consumer perceptions. In A. Cifuentes (Ed.), Comprehensive Foodomics (pp. 452-480). Elsevier. https://doi.org/10.1016/B978-0-08-100596-5.22876-X.

Statement of Authorship

Title of Paper	Authenticity and Traceability in the Wine Industry: From Analytical Chemistry to Consumer Perceptions
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Ranaweera, R. K. R., Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P. & Jeffery, D. W. (2021). Authenticity and Traceability in the Wine Industry: From Analytical Chemistry to Consumer Perceptions. Reference Module in Food Science. Elsevier. https://doi.org/10.1016/B978-0-08-100596-5.22876-X

Principal Author

Name of Principal Author (Candidate)	Ranaweera K. R. Ranaweera
Contribution to the Paper	Conceptualisation, investigation, visualisation, writing – original draft, writing – review and editing the final paper.
Overall percentage (%)	50
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.

Date 19/06/2020

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Lira Souza Gonzaga
Contribution to the Paper	Conceptualisation, investigation, visualisation, writing – original draft, writing – review and editing the final paper.

Date 22/06/2020

Name of Co-Author	Dimitra L. Capone
Contribution to the Paper	Conceptualisation, supervision, writing – review and editing.

Date 19/06/2020

Name of Co-Author	Susan E.P. Bastian
Contribution to the Paper	Conceptualisation, supervision, writing – review and editing.
Signature	Date 19/06/2020

Name of Co-Author	David W. Jeffery
Contribution to the Paper	Conceptualisation, project administration, resources, supervision, writing – review and editing the final paper.
Signature	Date 19/06/2020

Authenticity and Traceability in the Wine Industry: From Analytical Chemistry to Consumer Perceptions

Ranaweera KR Ranaweera[a], Lira Souza Gonzaga[a,b], Dimitra L Capone[a,b], Susan EP Bastian[a,b], and David W Jeffery[a,b],
[a] Department of Wine and Food Science, The University of Adelaide, Glen Osmond, SA, Australia; and [b] Australian Research Council Training Centre for Innovative Wine Production, The University of Adelaide, Glen Osmond, SA, Australia

Introduction 2
Wine as a Global Beverage 2
Wine Quality 2
Terroir 3
Wine Composition 3
Wine Authenticity 6
Overview of Authenticity 6
Applications of Wine Authentication 7
Implications of Wine Fraud 7
Authentication Methods for Wine 7
Omic Techniques for Wine 8
Analytical Techniques and Sensors 8
Mass Spectrometry 8
Separation and Detection Techniques 10
Sensor Technology 10
Spectroscopy 11
Chemometrics 13
Wine Traceability 15
Overview of Traceability 15
Application of Wine Traceability 15
Addressing the Issues of a Traceability System 15
Traceability and Authentication Assurance Systems 16
Tamper-Proofing Technologies 16
Covert Packaging Features 18
Blockchain Systems 18
Radio-Frequency Identification (RFID) Technology 20
Near Field Communication (NFC) 20
Barcodes Systems 20
Quick Response (QR) Codes 20
Wireless Sensor Network (WSN) 21
Consumer Perspective of Wine Authenticity 21
Authenticity Concepts and Dimensions 21
Looking for Authentic Products: The Consumer Profile 22
Branding Authenticity: What Do Consumers Want From an Authentic Brand? 22
Influence of a Wine Label on Authenticity Perception 23
Wine Authenticity and Innovation 23
Consumer Perspective on Traceability 24
Conclusion and Future Perspectives 24
Acknowledgments 25
References 25

Glossary

Ampelographic Associated with the identification and classification of grapevines by assessing anatomical features and morphological differences, such as the shape and color of leaves, and features of growing tips (hairless, shiny, etc.).
Chaosmetric Refers to technology that uses a non-human intervention approach to creating tags with unique features using a random and unreproducible process.
Fining agents Permitted wine additives that help protect wine against the development of physical or chemical instabilities.

Organoleptic Related to the senses, as in the perception of smell, taste, flavor, tactile sensations and appearance.
Terroir Delimited geographic location that links cultural traits and practices with the natural environment where a product is produced. Together these factors lead to originality, typicity and recognition of products from a region.

Introduction

Wine as a Global Beverage

Wine has been part of human history for thousands of years (McGovern et al., 2017) and endures today as a revered beverage in many cultures that combines potential health benefits with pleasure (Higgins and Llanos, 2015). Old World wines, represented by regions in Western and Southern Europe, and New World wines, from North and South America, South Africa, Australia, New Zealand, and more recently, others such as China, India and Japan, define the modern world of wine. Old World producers are recognized by their traditional and consistent geographical locations, viticultural practices, and winemaking techniques (Banks and Overton, 2010). On the other hand, New World producers are not as restricted by local viticultural and enological customs, and have benefited from innovations afforded through research and development coupled with interaction among different industry stakeholders (Aylward, 2003).

In relatively recent times, a noteworthy change has been observed in the international wine industry, with the perception that Old World producers no longer dictate the wine market as they once did (Aylward, 2003; Koutroupi et al., 2015; Campbell and Guibert, 2006; Wongprawmas and Spadoni, 2018). Reflecting this, wine production volume in Old World countries such as France and Italy has decreased between 1980 and 2016, whereas the opposite occurred for New World countries such as USA, Australia, New Zealand, Chile, and China. The same trend is true for domestic wine consumption in Old and New Worlds (Anderson et al., 2017; Koutroupi et al., 2015). The change is not only one of production but also of culture, and can be explained at least partially by consumers' growing appreciation of the traits that New World wines can provide, such as quality, uniqueness, and value for money (Aylward, 2003; Wongprawmas and Spadoni, 2018). Nonetheless, the changes also emphasize the truly global nature of wine production and consumption.

It is accepted that wines represent an association with the geographical location where they are produced, and consumers specifically demand trustworthy information about origin (Vergara et al., 2011), which it is often considered a significant factor in wine quality (Riovanto et al., 2011). Year of vintage, grape variety, and producer reputation are also viewed as major factors that can determine consumer preference and their notions of wine quality (Schamel and Anderson, 2003). When considering what consumers value, and within the context of a global market worth hundreds of billions of dollars, wine ultimately turns out to be no different to other consumer goods that need to be protected against various types of fraud. This requires the means to authenticate wines and verify product origin, variety and quality (Waterhouse et al., 2016a), and traceability systems to monitor the integrity of the grape and wine production processes.

This chapter provides up-to-date knowledge on authenticity and traceability in the wine industry. It covers the need for wine authentication, analytical methods and applications, the implications of wine fraud, and technological approaches to traceability that record information and help protect product integrity. Authenticity and traceability are also considered from the perspective of consumers, who are the ultimate stakeholders when it comes to the value of wine. The complexity of the topic requires some context, so the chapter begins by outlining the notion of wine quality, and explains its underpinnings in *terroir* and the importance of wine composition.

Wine Quality

Although it is a commonly used word, quality is a challenging concept to objectively describe and measure because of its broad meaning and close connection with hedonics (Hopfer and Heymann, 2014; Niimi et al., 2018). From a subjective standpoint, there are two ways to view the quality of wine. One option relies on the judgment of wine practitioners based on their assessment of a range of sensory characteristics, and especially a lack of faults (Charters and Pettigrew, 2007). A final wine quality score can be derived in this way, but difficulty arises due to the different approaches to such evaluations. Alternatively, wine quality can be judged from the consumer perspective, where there is reliance on the perceived quality, which can result from expected quality based on brand name, price and labeling, and experienced quality based on the actual sensory cues (Charters and Pettigrew, 2007). Perceived quality is a multidimensional concept involving a complex relationship between intrinsic (i.e., appearance, sensory traits, pleasure, origin, variety, typicity) and extrinsic (i.e., grapes, technically correct and consistent winemaking, marketing, packaging) elements (Charters and Pettigrew, 2007). Verdú et al. (2004) had previously suggested that red wine could be evaluated using a two-dimensional scale containing other dimensions (i.e., multidimensional): one dimension involved extrinsic attributes (region of origin, image/renown, and presentation/packaging), and the other related to intrinsic attributes (age, harvest information, sensitivity (balance/body), and acuteness (intensity/complexity of bouquet)).

D'Alessandro and Pecotich (2013) noted that assessment of sensory quality can be a difficult task for inexperienced consumers, who rely mainly on country of origin and brand name (to a lesser extent) as indicators of quality, whereas experts were able to

objectively use sensory quality in association with price, brand and country of origin. Hopfer and Heymann (2014) demonstrated that consumers could evaluate quality through tasting experiences just as well as an expert panel, with the only limitation being the difficulty of clearly communicating their opinions. Nonetheless, consistently judging wines can be challenging even for experts due to the highly variable matrix and dimensions involved in wine quality (Hopfer and Heymann, 2014).

The complexity of wine aromas and flavors means that no one chemical compound or sensory descriptor is expected to completely explain all the aspects involved in a wine quality definition (Hopfer et al., 2015). Instead, the resultant sensory attributes are reliant upon synergies and interactions between a broad range of volatile and non-volatile compounds (Hjelmeland et al., 2012). As an example, analysis of Cabernet Sauvignon wines by Hjelmeland et al. (2012) showed a large number of chemical constituents could be positively as well as negatively correlated with aroma and flavors attributes of the wine, which suggested the existence of additive and interactive effects among compounds in the wine matrix. Despite the impact of the matrix, certain sensory attributes can be considered as negative or positive in terms of quality, or even both, depending on the intensity of those characters in the wine (Hopfer et al., 2015).

Even though sensory and chemical attributes can be correlated with hedonic aspects, quality can be a broad concept and can be represented through a number of factors for the consumer. For many, the region of origin is a determinant for a purchase decision, and can be the most important information used to predict quality by raising a quality expectation of the region (Ray and Johan, 2007). For consumers with high involvement in the wine sector, wine that expresses the typicity expected of the region is an indicator of high quality (Tustin and Lockshin, 2001; Charters and Pettigrew, 2003). Just as regionality can be considered an aspect of wine quality, a high average quality can boost a wine region's reputation and shape the regional individuality (Easingwood et al., 2011). Wine typicity and regionality have their origins in the environment where the grapes are grown, which falls within the broad concept of terroir.

Terroir

The definition of terroir is important because it relates to the variables that need to be considered when studying authenticity. Terroir is a French word that cannot be precisely translated into English and can have a broad spectrum of definitions when described in English depending on the perspective of the writer (Gladstones, 2011). It is a concept that is discussed worldwide and can be used for different contexts. During a UNESCO gathering, professionals from two French institutes (INRA/INAO)[1] proposed a definition for terroir as:

> Terroir is a delimited geographic space, defined by a human community that during its history builds a distinct set of cultural traits, knowledge and practices, based on a system of interactions between the natural environment and human factors. This applied knowledge creates an originality, confers a typicity and allows products and services from this space to be recognized.
>
> (translated and adapted) (UNESCO and Association Terroirs & Cultures, 2005; Unwin, 2012)

According to Unwin (2012), the definition of terroir covers a broad range of factors and goes beyond physical dimensions that some notions are restricted to. In order to be scientifically valid, however, terroir has to be given an objective and specific meaning, with studies dedicated to terroir primarily interested in understanding the impacts on grape and wine quality of environmental aspects in vineyards or specific regions (Vaudour, 2002). Terroir also relates to the authenticity of a product by defining the factors, such as climate and soil conditions that directly affect the chemical composition of grapes and ultimately translate into differences in wine composition. Winemaking inputs such as fermentation conditions, yeast strains, and processing steps will further affect wine constituents (Arvanitoyannis et al., 1999), and can also be considered as important aspects of terroir.

Wine Composition

Investigating wine authenticity first requires knowledge of wine composition (**Fig. 1**), which is a complex fermented beverage that mainly consists of water and ethanol (Waterhouse et al., 2016b). Accounting for about 97% w/w of the constituents in wine, ethanol and water are mainly responsible for the physicochemical properties of wine, including viscosity, polarity and solubility, but also have direct and indirect effects on organoleptic properties (Margalit, 2004; Waterhouse et al., 2016c). The remaining 3% of components represent the main contributors to the perceived color, aroma and flavor of the wine, and other sensory attributes that drive style, quality and consumer liking. These components can be categorized into non-volatile and volatile compounds. The non-volatiles primarily include phenolic compounds, sugars, glycerol, proteins, amino acids, organic acids and inorganic compounds (Alañón et al., 2015). The volatile compounds include higher alcohols, terpenes, esters, aldehydes, ketones, volatile acids, and heteroatomic compounds involving sulfur and nitrogen (Villamor and Ross, 2013). Brief examples provided in the following paragraphs indicate how such components have been studied for authentication purposes (see Section "**Authentication Methods for Wine**" for detailed information about analytical approaches).

[1]INRA – Institut National de Recherche Agronomique en France/INAO – Institut National des Appelletions D'origine.

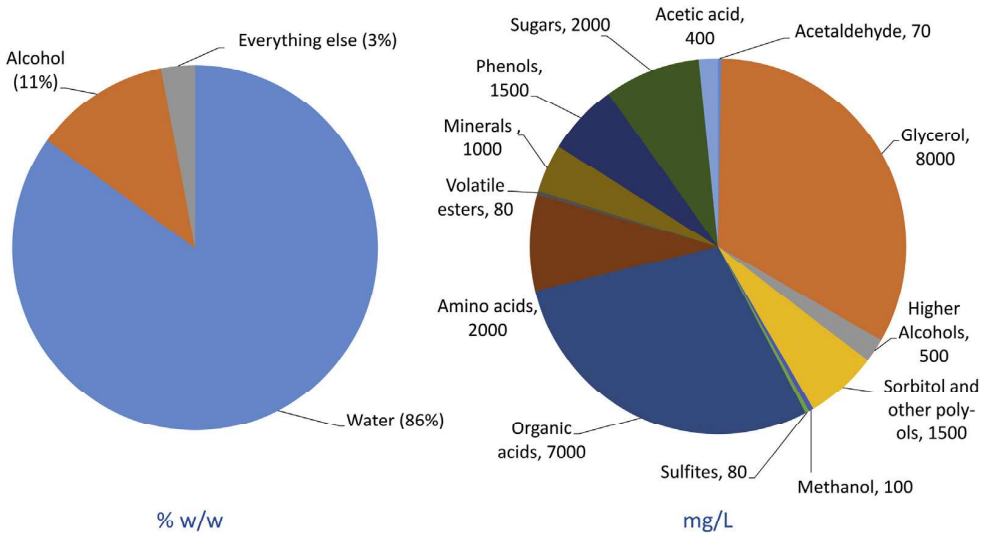


Figure 1 Chemical composition of dry red table wine with major components given on a weight-for-weight (w/w) % basis, and "Everything else" presented in mgL^{-1}. Adapted from Waterhouse et al. (2016b).

Grape carbohydrates, mainly consisting of the monosaccharides glucose and fructose, play a major role as the carbon source for yeast during fermentation. They can also exist in polymerized forms, as pectins and other polysaccharides, and glucose and certain monosaccharides form glycosides with anthocyanidins (grape pigments) and other molecules (Jackson, 2014). Residual sugars, for example arabinose and xylose, are also present but are less than 1.5 gL^{-1} and do not contribute sweetness to wine (Jackson, 2014). Aside from ethanol as the major product of fermentation, glycerol is also derived from glycolysis and is a relatively abundant sugar alcohol in wine, where it may contribute to perceptions of wine body and mouthfeel (Zhao et al., 2015). The $\delta^{13}C$ isotopic ratio of glycerol, ethanol and grape sugars has been applied in wine authentication, providing information on sugar origin (Guyon et al., 2011). Other polyols constitute additional sugar alcohols but are present to a much lesser extent than glycerol.

Organic acids are another important class of constituents in wine, with some being formed in grapes (tartaric, malic, and citric) and others formed by yeast or bacteria during winemaking (succinic, lactic and acetic) (Fig. 2). Tartaric acid is the most prominent acid in ripe grapes and is mainly responsible for the pH of wine but others, including an array of volatile acids (e.g., acetic), contribute to total acidity (Waterhouse et al., 2016d). Even though organic acid profile is not commonly used as a characteristic index for wine authentication, it has been applied successfully (Huang et al., 2017).

The mineral content of wine arises in diverse ways, including from irrigation water, vineyard soils, processing aids, winery equipment and external contaminants. The concentrations and oxidation states of these metals not only affect wine sensory properties and stability, but their profiles can also be used for authentication purposes (Waterhouse et al., 2016e).

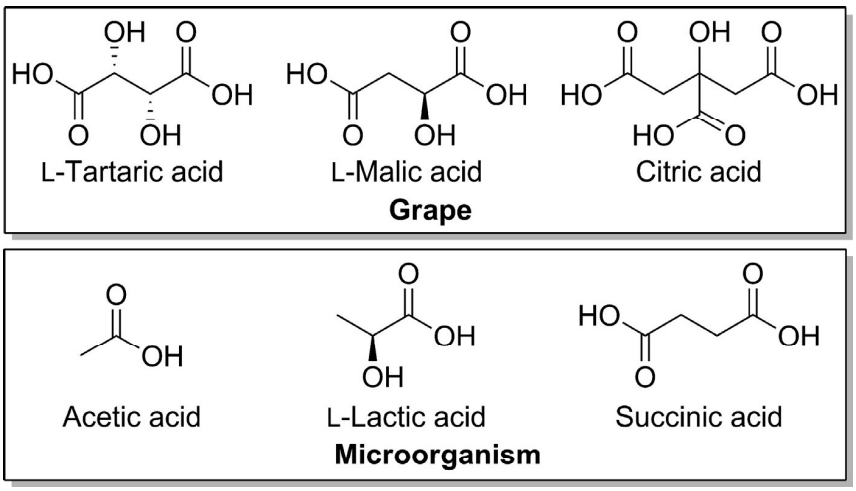


Figure 2 Major organic acids present in wine from grape and microbiological sources.

Phenolic compounds are one of the most significant components of wine responsible for organoleptic properties such as color, astringency, and bitterness; they are also important from a health perspective as bioactives with antioxidant and anti-inflammatory properties (Waterhouse, 2002). Phenolics are found at various concentrations depending on the type (red or white) and variety of grape (Waterhouse et al., 2016f), and their profiles can be used successfully for varietal authentication. Many of the phenolics are found in grape seed and skin, with some occurring in the flesh. They are extracted during wine production, especially for red wine-making where maceration with grape solids occurs. Wine phenolics can be categorized into two groups according to the carbon structure: as flavonoids (C_6–C_3–C_6) and non-flavonoids (**Fig. 3**).

Flavan-3-ols, flavonols, and anthocyanins are the main flavonoids, with phenolics such as hydroxycinnamates, stilbenes, and benzoic acids classified as non-flavonoids (Waterhouse et al., 2016g). White wine has a lower concentration of phenolic compounds and a higher abundance of hydroxycinnamoyl tartaric acid esters, whereas red wine is characterized by having anthocyanin pigments and tannins (Garrido and Borges, 2013).

Nitrogenous compounds present in grapes and wine are of special interest due to their effects on fermentation performance (as a nutrient) and role in protein haze due to heat instability (Margalit, 2004). The content of amino acids in finished wine tends to be low but they are important as a source of volatile higher alcohols produced during fermentation (Huang and Ough, 1991) and essential in decreasing H_2S formation (Mendes-Ferreira et al., 2011). The concentrations of the major amino acids present in wine, including proline, arginine and glutamine (**Fig. 4**), vary among different varieties and can be used as a marker for varietal authentication of wine (Herbert et al., 2000).

Volatiles that can contribute to the aroma profile of wines include aldehydes and ketones, with acetaldehyde being the most abundant. Esters are another class of volatiles that are formed primarily during fermentation and are significant due to their *organoleptic* contributions, especially of pleasant fruity aromas (Villamor and Ross, 2013). Although a range of trace volatiles are absent from **Fig. 1** due to their low concentrations, some are important as character impact compounds in certain varietal wines. Examples include rotundone (a sesquiterpene that contributes pepper aroma in Shiraz), polyfunctional thiols (citrus and tropical fruit aromas in Sauvignon blanc, such as 3-sulfanylhexan-1-ol, 3-SH) and methoxypyrazines (green capsicum and vegetal aromas in Cabernet Sauvignon, such as 3-isobutyl-2-methoxypyrazine, IBMP) (Jeffery and Wilkinson, 2014; **Fig. 5**). Volatiles are one of the major groups of compounds that can be utilized for wine varietal authentication (Villano et al., 2017).

Overall, given the links between region of origin, vineyard, winery, and the compositions of fruit and wine, characterization of the aforementioned chemical markers is critical when attempting to determine the identity and authenticity of a specific wine. Considering the broader supply and distribution chains and the use of various processing aids, bottling equipment, vineyard blocks, etc., traceability also becomes an important additional element that is worthy of attention. These aspects are introduced in the following sections.

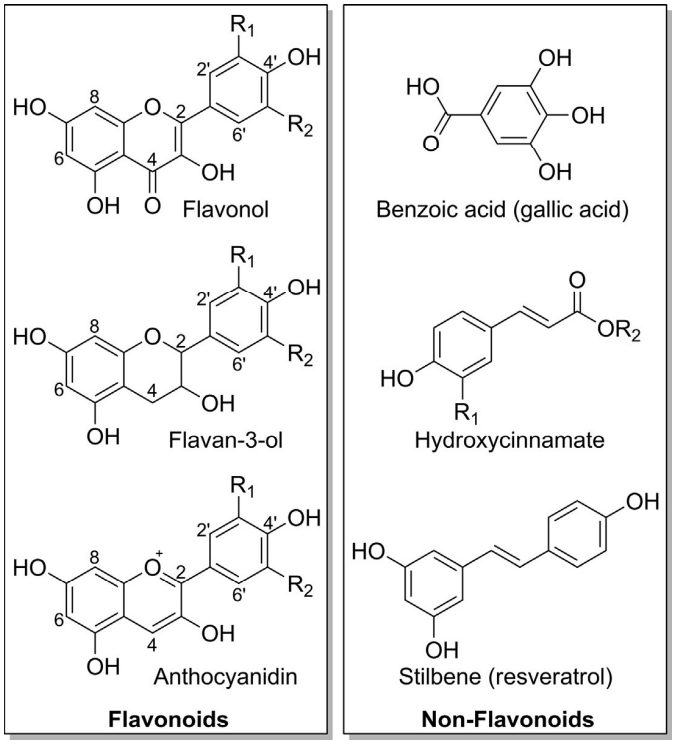


Figure 3 Examples of phenolic compounds present in wine that are classified as flavonoids and non-flavonoids.

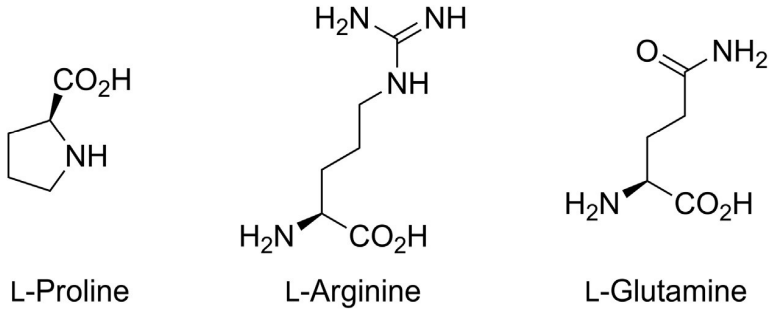


Figure 4 Major amino acids present in wine.

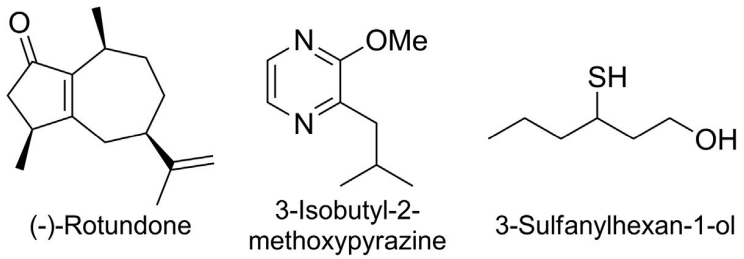


Figure 5 Trace volatiles found in wine include rotundone (left), methoxypyrazines such as IBMP (center), and polyfunctional thiols such as 3-SH (right).

Wine Authenticity

Overview of Authenticity

Ensuring the identity of a food product has become more of a concern in recent times, especially with the advance of globalization, which invokes a borderless economy that facilitates easier transportation of food commodities across international borders (Carcea et al., 2009). Hence, it has become a challenge for government authorities and industry stakeholders to confirm the identity and provenance of food products throughout the food supply chain (Aung and Chang, 2014). Another reason identified for the increased interest in food authentication is the rise in food fraud, where the product has been altered for economic gain (Pustjens et al., 2016). Most importantly, any failure in quality assurance of food not only has serious economic consequences but also bears a potential risk for human (and animal) health (FAO/IAEA, 2017).

In general, food authenticity can be defined as the process that certifies the product is the same as described on its label (Danezis et al., 2016). However, criteria that define the authenticity of foods are numerous and vary from product to product, and include geographical origin, botanical origin, composition, year of manufacture, and production process (Pascu, 2013). Therefore, the detection and investigation of potentially fraudulent products or processes requires the implementation of various methods and techniques appropriate to the food product in question.

In the particular case of wine, authenticity is a paramount factor when considering quality and consumer confidence in the product. Details present on the wine label, such as brand, variety, origin and vintage, are defining characters associated with wine style and quality that drive consumer expectations (Palade and Popa, 2014). Besides, the complex chemical nature, high prices commanded, and availability around the world, mean that wine is more vulnerable to adulteration than other foods. Hence, guaranteeing the quality and correct information on wine labels, through strict guidelines imposed by authorities, analytical methods, and traceability systems, is important for the wine industry (Makris et al., 2006).

Historically, sensory evaluation has been used to assess wine quality and authenticate wine products, and it still serves that purpose. However, this approach is rather subjective and may lead to misinterpretation, so objective approaches such as those based on analytical methods have been introduced and applied (Kallithraka et al., 2001). Markers of geographical origin, variety, and vintage are considered as the main identity-defining traits that help to determine the value of a wine and therefore have to be considered when developing methods for authentication of wine (Médina et al., 2013). Primarily, these traits can be assessed using chemical markers such as phenolic and volatile compounds, elemental composition, stable isotope ratios, and amino acid profiles (Vergara et al., 2011; Kokkinofta et al., 2017). Given the complexity of the datasets, multivariate data analysis methods (i.e., chemometrics) are often utilized to identify the patterns or classify groups of particular wines (Geana et al., 2016a). Overall, the chosen approach is important in verifying the authenticity of wine and confirming that a wine is in compliance with its label description (Schlesier et al., 2009).

Applications of Wine Authentication

More evidently in recent times, popular wines of the world are every so often subject to adulteration, substitution or counterfeiting (Waterhouse et al., 2016a). Therefore, authentication of wine has become inevitable to address such fraudulent activities. Wine authentication gives an assurance to the consumer that they are buying an original product and getting what they paid for, but also encompasses aspects of food safety in the case of adulteration. One of the oldest incidents reported is the addition of lead acetate to increase wine sweetness, which caused severe health effects due to lead poisoning (Lachenmeier, 2016). Another well-known example involved the addition of diethylene glycol to enhance the quality of wine and mask the addition of sugar (Holmberg, 2010), but diethylene glycol is not a legal wine additive due to its toxicity (Sebastian, 2009). Similarly worrying, addition of exogenous methanol to increase alcohol content has caused fatalities in Italy (Suro, 1986). In addition to safety assurance, authentication provides a marketing advantage for producers who can protect their brands among unauthenticated wines.

Because a particular geographical area where a product arises is important to perceived quality and style, a main focus of wine authentication is confirming the origin of wine. EU regulation No 1308/2013 specifies Protected Designation of Origin (PDO) and Geographical Indication (GI), which are terms used to define and protect wines from specific geographical areas. These regulations have significant implications for the reputation, integrity and price of wine, and being able to confirm the origin of wine bestows the advantage of differentiating it according to the region. For example, findings from a study of the wine industry in South Africa suggested that use of GIs increased the value of South Africa's wine exports into the EU (Lubinga et al., 2017).

Other than geographical origin, year of vintage and grape variety (cultivar) are also considered as major factors that determine the value of a wine. Compositional and sensory elements are largely dependent on variety, underlining the importance of varietal authentication (Rapeanu et al., 2009). In the same vein, wine age defines the physical, chemical and biochemical characteristics of wine, and the year of vintage can add great value to the overall sensory characteristics in years with favorable grape growing conditions (Danezis et al., 2016). Therefore, development of particular methods for authentication of wine based on a range of intrinsic factors has been of growing interest to all stakeholders of the wine sector.

Another application of wine authentication is counteracting wine fraud. This aspect includes identifying different means of adulteration, such as dilution, unapproved enhancements like sweeteners or taste/mouthfeel enhancers, and substitution, with less expensive varieties or regions (Waterhouse et al., 2016a). Moreover, authentication techniques enable identification of false statements regarding production processes. For example, in the authentication of rosé wine, identifying the source of raw material is important, whether it is red grapes, a mixture of red and white grapes, or addition of extracts from grape pomace, that yield the color (Rapeanu et al., 2009). Similarly, in sparkling wine production, identifying the origin of CO_2 (endogenous or exogenous incorporation) is required because this dictates quality and style. Notably, prestigious wines are often subject to adulteration (Waterhouse et al., 2016a), so identifying and protecting the reputation and integrity of a product/brand against fraud is an important challenge for modern analytical chemistry.

Implications of Wine Fraud

Wine fraud can be viewed in terms of extrinsic and intrinsic factors. Falsification of cultivar and geographic origin, mislabeling, and replacement with low quality wine are considered as extrinsic manipulations whereas modifying original components, such as through addition of water, ethanol, sugar or other substances for coloring and flavoring, are reflective of intrinsic processes (Holmberg, 2010). Wine laws include information on permitted additives and wine production practices and were introduced for various purposes, but mainly for consumer health and safety. Laws are also designed to overcome wine fraud by verifying the origin of wine through regulated protected designations of origin and correct labeling (Fandl, 2018). Along with strict regulations comes the need for authenticity testing and a secure traceability system to guarantee the authenticity of wine.

The economic impact of wine fraud can be substantial for any country, region, or producer engaged in wine trade, and various instances have had potentially serious financial consequences. For example, the so-called Brunello wine crisis in Italy is one where producers were indicted for not adhering to production standards, having used a different grape variety than permitted (Cavicchi and Santini, 2011). Another incident involved the replacement of Pinor Noir wine with cheaper wine that was exported to the USA as Pinor Noir from France (Styles, 2009). Perhaps the most infamous example of wine fraud is that of Rudy Kurniawan, who sold millions of dollars' worth of fake wine before he was caught and who featured in the 2016 documentary "Sour Grapes" (Hellman, 2008).

Authentication Methods for Wine

The increase in counterfeit products and the growing consumer interest in verifying product origin and quality have added to the drive for methods that provide information on authenticity. Traditional analytical tools are still valuable assets for authentication, but with the advancement of rapid technologies such as biological/chemical fingerprinting, a range of other product validation methods have been developed. Specifically, high-throughput technologies termed "omics" have gained attention by introducing various measurements from gene to metabolite level that are categorized as genomics, transcriptomics, proteomics, and metabolomics (Sébédio and Malpuech-Brugère, 2016). Application of these techniques in determining the adulteration and authentication of wine has been a powerful approach in recent times (Pinu, 2018).

Omic Techniques for Wine

The genotypic and phenotypic traits of grapevine, yeast and bacteria shape the composition of finished wine, giving it specific measurable characteristics. Genomics permits identification of the genetic variations and uniqueness of a particular wine (Borneman et al., 2013). Application of DNA markers for varietal authentication of wine has been advantageous over traditional *ampelographic* methods because morphological characteristics can be affected by diseases, vine growth stage, and different environmental conditions (Işçi et al., 2009). Moreover, the stability of DNA under various wine processing conditions such as high temperature or low pH makes it suitable as a marker for wine authentication (Villano et al., 2017). Identification of a correct DNA sequence is vital in authentication methods and a couple of methods have been applied: simple sequence repeat (SSR) and single nucleotide polymorphism (SNP). Large databases (European Vitis database; Swiss microsatellite database) containing genetic profiles of wines have been developed using these sequencing methods. In particular, the International Organisation of Vine and Wine (OIV) has approved the SSR method for varietal authentication. Similarly, SNPs are being widely applied in the grape and wine industry to identify first-degree genetic relationships of grapevine accessions (Barrias et al., 2019).

Transcriptomics is not directly utilized in wine authentication, but it has been applied to the analysis of variations in different yeast strains from diverse origins (Mendes et al., 2017). A transcriptome reflects the expression of mRNA at any given time and conditions, thus determining the pattern of gene expression and likely changes in organism response and metabolite production. It is important in analyzing the influence of yeast strains on wine flavor and fermentation bouquet, for example. Transcriptomic techniques like next-generation sequencing (RNA-seq) are essential for genomic studies as they facilitate the understanding of the functional elements of the genome (Van Emon, 2016).

Proteomic-based techniques that consider the proteins present in wine are another emerging area for authentication purposes. Proteomics has been used in discrimination of white wines according to the variety (Rešetar et al., 2016) and for detection of adulteration based on proteins from *fining agents* (Cereda et al., 2010). However, these techniques are not widely applied mainly due to the trace levels of proteins remaining in finished wine, and the presence of exogenous proteins from yeast or bacteria that pose further difficulties when attempting to evaluate aspects such as origin, vintage, and variety (Ortea et al., 2016).

Among the "omic" techniques, metabolomic approaches that involve comprehensive analysis of metabolites are the most widely applied in wine authentication. These techniques can be categorized into targeted and non-targeted methods, depending on the objective of the analysis. Specific marker compounds/metabolites are considered in targeted analyses and variations in concentrations are used for differentiation of samples. On the contrary, non-targeted analyses such as by spectroscopic techniques and sensor technologies (e-nose and e-tongue) provide a chemical fingerprint of the sample, with similarities/differences in fingerprint being considered for classification. Other than the main omics techniques described above, isotope profile analysis (of elements, water, carbon) – as in isotopolomics – is also established for wine authentication. It is mainly applied in geographical authentication because of the relationship between the soil profile where grapes are grown with the elemental profile of wine (Coetzee et al., 2014). Recently, various analytical techniques based on metabolomics or isotopolomics have been developed, resulting in the detection of additional identity markers. Along with improvements in data analysis from various instruments or multiple sources, the analytical methods now available have enhanced the possibilities for wine authentication (Versari et al., 2014).

Analytical Techniques and Sensors

Different analytical methods applied in metabolomic and isotopolomic studies for wine authentication are useful in various ways, such as to: (a) identify and quantify metabolites; (b) discriminate samples in a population; and (c) to predict categories for building statistical models (Cevallos-Cevallos et al., 2009). Often, using a combination of both qualitative and quantitative data blocks from different methods (targeted or non-targeted) and data fusion are important in determining authenticity by providing more dimensions of information about the sample (Borràs et al., 2015). Furthermore, sample complexity and the different magnitudes of analyte concentrations mean that it is useful to couple different methods, such as separation techniques with detection by spectrometry or spectroscopy, although rapid and simple methods would tend to be the most ideal for implementation by industry. The following section describes the main analytical methods used in wine authentication.

Mass Spectrometry

In mass spectrometry (MS), ionized metabolites are resolved based on their mass-to-charge (m/z) ratio (Rubert et al., 2015). MS has been used as a standalone tool for varietal, geographical, and vintage authentication (in conjunction with chemometrics) (**Table 1**) as well as being coupled with separation techniques such as gas or liquid chromatography (**Table 2**) and chemical sensor systems such as e-nose (MS e-nose) for detection and quantification of chemical constituents (covered in Sections "**Separation and Detection Techniques**" and "**Sensor Technology**"). However, in terms of practical applicability within an industry setting, some methods may rely too heavily on time and labor, and utilize complex instrumentation that is not readily available. Therefore, application of techniques that use an ambient ion source like direct analysis in real time (DART)-MS have been introduced (Guo et al., 2017). A study by Rubert et al. (2014) compared different authentication strategies involving high resolution MS with quadrupole time-of-flight (Q-TOF-MS) coupled with DART (**Table 1**) or ultra-high performance liquid chromatography (UHPLC), using wine samples that were measured directly as well as a polyphenol fraction isolated from wines. They concluded that DART was suited to a wide range of analytes and offered greater efficiency but lacked the discrimination power afforded by UHPLC-Q-TOF-MS.

Table 1 Examples of mass spectrometry approaches used for wine authentication

Technique	*Chemical marker*	*Classification method*[a]	*Parameters for authentication*	*References*
DART–HRMS	Flavonol glucosides and polyphenols	OPLS-DA	Varietal differentiation	Rubert et al. (2014)
MALDI-TOF-MS	Peptide profile	PCA HCA	Varietal and processing method	Rešetar et al. (2016)
PTR-MS	Volatile profile	ANOVA PCA	Vineyard discrimination	Schueuermann et al. (2017)
ICP-MS	Li, Na, Mg, Si, P, K, Ca, Mn, Fe, Ni, Zn, Rb, Sr, Cs, Ba	LDA (up to 100% classification)	Geographical origin of Australia	Martin et al. (2012)
ICP-MS	Tl, U, Li, Rb, Mg	CA (100% classification)	Geographical origin of Argentina, Brazil, Chile, and Uruguay	Bentlin et al. (2011)
IRMS	$^{13}C/^{12}C$ and $^{18}O/^{16}O$	–	Geographical origin and vintage	Magdas et al. (2012)
IRMS	$\delta^{13}C$ and $\delta^{18}O$	ANOVA, LDA	Adulteration by addition of sugar and water	Geana, Popescu, Costinel et al. (2016b)
TIMS and IRMS	$^{87}Sr/^{86}Sr$ and $\delta\ ^{13}C$	GPA	Terroir differentiation	Di Paola-Naranjo, Baroni, Podio et al. (2011)
MS e-nose	Volatile profile	LDA, SLDA	Geographical origin of Sauvignon Blanc	Berna et al. (2009)
MS e-nose	Volatile profile	PCA, LDA, DPLS	Varietal differentiation of Riesling and Chardonnay	Cozzolino et al. (2005)

[a]See Section 3.6 of this chapter for details.

Table 2 Examples of separation and detection techniques used for wine authentication

Technique	*Chemical marker*	*Classification method*[a]	*Parameters for authentication*	*References*
HPLC-PDA	Targeted (anthocyanins)	PCA, CDA	Varietal discrimination of Blaufrankisch (Lemberger), Saint Laurent and Blauer Portugieser	Kumšta et al. (2014)
HPLC-PDA	Targeted (benzoic acids and stilbenes)	DA	Varietal and regional discrimination	Kallithraka et al. (2007)
HPLC-MS/MS	Targeted (phenolics)	CDA	Region, varieties and vintage	Jaitz et al. (2010)
UHPLC-Q-TOF-MS	Non-targeted metabolite fingerprinting (polyphenols)	PCA, OPLS-DA, SIMCA	Varietal, geographical and vintage	Rubert et al. (2014)
HPLC–Q-TOF-MS	Non-targeted metabolite profiling (anthocyanins)	PCA, PLS-DA	Varietal discrimination	Vaclavik et al. (2011)
UPLC-Q-TOF-MS, FTICR-MS	Non-targeted metabolite profiling (caftaric acid, rosmarinic acid, caffeic acid, and aesculetin)	PCA	Geographical origin	Roullier-Gall et al. (2014)
HS-SPME-GC-MS	Targeted (esters)	PCA, SLDA	Varietal differentiation of Cabernet Sauvignon, Cabernet Gernischt and Merlot	Zhang et al. (2010)
HS-SPME-GC-FID	Non targeted Volatile compounds (ethyl acetate, isoamyl alcohol, ethyl lactate)	PCA, CA, LDA	Varietal differentiation	Márquez et al. (2008)
HS-SPME GC × GC-TOF-MS	Non targeted Volatile compounds (2,3-butanediol, 4-carene, 3-penten-2-one, diethyl succinate)	PCA, SLDA	Differentiation of Cabernet Sauvignon, Merlot, Chardonnay, Sauvignon Blanc and Pinot Noir varieties.	Welke et al. (2013)
HS-SPME GC × GC-TOF-MS	Non targeted Volatile compounds (3-isobutyl-2-methoxypyrazine, menthone, isomenthone)	CVA and PLS-DA	Differentiation of Cabernet Sauvignon from ten different Australian regions	Robinson et al. (2012)

[a]See Section 3.6 of this chapter for details.

Matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS) is another atmospheric pressure ionization technique used with MS in the characterization of thermally labile molecules such as proteins (Dreisewerd, 2014). It is considered a simple but sensitive and fast method compared to chromatographic approaches (Danezis et al., 2016). Laser energy is used to vaporize and ionize analytes mixed within a matrix, and ions formed are analyzed by MS. This approach has been applied to

the classification of Spanish (Nunes-Miranda et al., 2012) and Croatian (Rešetar et al., 2016) white wines on a varietal basis using wine peptide profiles. According to Rešetar et al. (2016), white wines produced by different production techniques could be differentiated, such as filtered vs non-filtered, which could be useful in production process control. However, further development of the protocol in identifying authenticity markers for regional and vintage classification would be necessary.

Proton transfer reaction mass spectrometry (PTR-MS) is another variation of MS that can be employed (in real-time) for quantitative analysis of volatile organic compounds. This technique uses chemical ionization of water vapor to yield hydronium ions (H_3O^+) that react with the analytes to produce protonated ions for MS detection. However, a disadvantage in wine analysis is that ethanol interferes with the proton transfer reaction (Schueuermann et al., 2017; Sémon et al., 2018). Several studies have used different methods to overcome ethanol interference in wine, such as dilution of the headspace ethanol concentration with high purity nitrogen (Spitaler et al., 2007) or separation of ethanol from the wines by using fast GC combined with PTR-TOF-MS (Romano et al., 2014).

Inductively coupled plasma (ICP)-MS is used for elemental analysis and can rapidly quantify multiple elements (especially metals) at ultra-trace to trace levels ($<1\ \mu gL^{-1}$ to mgL^{-1}). In ICP-MS analysis, the analytes in the sample are decomposed to elements and ionized by a high-temperature argon plasma prior to detection. It is more advantageous than other methods used to measure elements, such as atomic absorption spectroscopy (AAS), because of the lower detection limits, speed, and multi-element capabilities (Luykx and van Ruth, 2008). ICP-MS has been used for geographical origin determination across countries and regional discrimination within a country (Martin et al., 2012; Bentlin et al., 2011; **Table 1**).

For authentication of wine in isotopolomic studies, the commonly used spectrometric methods are isotope ratio MS (IRMS) and thermal ionization mass spectrometry (TIMS), which are based on combustion or vaporization and ionization of samples to enable precise MS detection of different naturally-occurring isotopes (**Table 1**). Light isotopes are analyzed by IRMS whereas heavier elements utilize TIMS. These methods have been used (with statistical treatment of the data) to determine the ratios of carbon, oxygen, hydrogen or strontium isotopes for geographical origin determination (Di Paola-Naranjo et al., 2011; Magdas et al., 2012) and identification of adulteration (Geana et al., 2016b; **Table 1**). Additionally, IRMS can be hyphenated with separation techniques that enhance the analysis of different isotope ratios (Christoph et al., 2015). The OIV, of which many wine producing nations are members, specifies IRMS as an official method for wine and must analysis, for authentication or analysis of metals (Christoph et al., 2015).

Separation and Detection Techniques

Chromatographic methods are the most commonly applied separation technique in food analysis. Liquid chromatography (LC) is the technique used for non-volatile liquid samples, as in the analysis of water soluble substances like anthocyanins (Košir et al., 2004). Advanced types of LC methods, such as high performance LC (HPLC) or ultra-high performance LC (UHPLC) have increased the speed, resolution and sensitivity of the application. Various detectors used with LC, such as ultraviolet (UV) and photo diode array (PDA) (**Table 2**), are useful for targeted analysis of phenolic compounds (Jaitz et al., 2010) and fatty acids (Della Corte et al., 2013), for example. However, it is important to select the most appropriate detector to use for authentication as each have their own drawbacks, such as the low resolution of UV detectors (Liu et al., 2017) and lack of discrimination by PDA detectors. Nonetheless, HPLC with PDA has been applied to the analysis of phenolics as a simple and less-expensive technique compared to other detectors. On the other hand, coupling LC separation (which may still include UV or PDA detection) with MS is a very effective solution for characterizing a range of wine constituents (Di Stefano et al., 2012) including the targeting of phenolics (**Table 2**). A number of studies have applied TOF-MS to untargeted metabolite profiling (Vaclavik et al., 2011; Rubert et al., 2014; Jaitz et al., 2010). This extends to the combination of UPLC-Q-TOF-MS with in silico alignment of exact mass data from Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS), which brings improvements by increasing resolution and mass accuracy for analyte identification. This approach has revealed a diverse range of metabolites and provides greater opportunity to classify samples into groups (e.g., by geographic origin irrespective of vintage) based on a comparison of marker compounds (Roullier-Gall et al., 2015).

As a complementary technique to LC approaches, gas chromatography (GC) is used to separate volatile or semi-volatile compounds and has been applied to the study of authenticity based on the volatiles present in wine. For determination of these compounds, it is often necessary to apply sample extraction and preparation techniques such as liquid-liquid extraction (LLE), solid phase extraction (SPE) or solid phase micro extraction (SPME) prior to instrumental analysis (Vilanova and Oliveira, 2012). As with LC, different detectors are coupled to GC to obtain quantitative measurements. In wine analysis, flame ionization detector (FID) and MS are frequently used for targeted or non-targeted analyses and varietal differentiation (**Table 2**) but other detectors are also common (e.g., selective detectors for sulfur or nitrogen). GC–MS is advantageous because of the ability for simultaneous determination of different metabolite classes, coupled with high chromatographic resolution (Pinu, 2018). Comprehensive two-dimensional gas chromatography (GC × GC), involving two GC columns with orthogonal stationery phase chemistry, provides a more powerful analytical method compared to a single column GC, and offers high peak capacity, selectivity and sensitivity for wine authentication (Welke et al., 2013). Varietal differentiation (Welke et al., 2013) and geographical authentication (Robinson et al., 2012) using volatile markers has been accomplished using GC × GC (**Table 2**).

Sensor Technology

Chemical sensor systems used for wine analysis, namely electronic nose (e-nose) and electronic tongue (e-tongue), are developed to mimic human olfaction (Śliwińska et al., 2016; **Fig. 6**). The process involves capturing a fingerprint of the flavor or aroma molecules

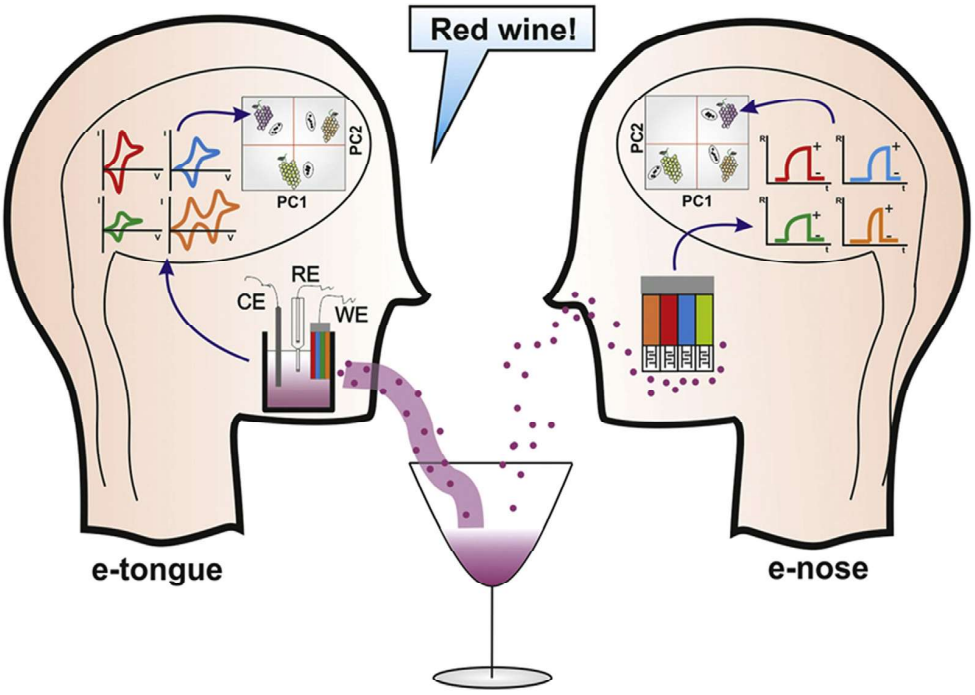


Figure 6 Scheme of the working principles of e-nose and e-tongue. Reproduced from Rodríguez-Méndez, M.L., De Saja, J.A., González-Antón, R., García-Hernández, C., Medina-Plaza, C., García-Cabezón, C., and Martín-Pedrosa, F., 2016. Electronic noses and tongues in wine industry. Front. Bioeng. Biotechnol. 4:81. Copyright 2016. License at https://creativecommons.org/licenses/by/4.0/.

from a sample with a sensor (often a non-specific array but MS can be used) and using multivariate data analysis to distinguish the specific odor or taste (Mannino et al., 2007).

Favored for its minimal sample preparation and rapid analysis, a number of studies have utilized e-nose for authentication (Cynkar et al., 2010; Berna et al., 2009). In the application of an e-nose for wine discrimination, the extraction of volatiles is firstly accomplished using techniques such as static headspace (HS), purge and trap (P&T) or SPME (Lozano et al., 2016). This is an important step because the headspace composition needs to represent the wine's original aroma profile. An array of gas sensors then transform the chemical signals obtained from wine volatiles into electronic signals. Different sensors such as metal oxides (MOX) (Berna et al., 2009), conducting polymers (CP), or surface acoustic wave sensors (SAW) (McKellar et al., 2005) have been used (Guadarrama et al., 2001) for authentication of wine by variety and geographically. MS-based e-nose has also been applied (see **Table 1**) and has the advantage of overcoming any interference from ethanol when analyzing headspace constituents of wine (Cozzolino et al., 2005). Finally, a pattern recognizing system classifies the wines according to specific categories by supervised learning or other multivariate types of analysis (Cozzolino et al., 2005).

Spectroscopy

Various spectroscopic techniques are available for wine fingerprinting, including near infrared (NIR), mid infrared (MIR), Raman, ultraviolet–visible (UV–vis), and nuclear magnetic resonance (NMR). Based on the electromagnetic (EM) spectrum, the interactions of EM radiation with molecules, compounds, atoms or nuclei, are recorded and represented in the form of spectra, which depend on the properties of the analyzed sample and the wavelength range (**Fig. 7**). Application of appropriate chemometric methods are vital with spectroscopic analyses, and there are several data preprocessing steps to correct for baseline drifts, scattering and path length variation (Oliveri and Downey, 2013).

Being robust and reproducible, NMR spectroscopy is a widely-used technique that requires minimal pre-processing of samples (if doing direct measurements of wine), and has the added advantage that the inherent properties of the sample are maintained (Alañón et al., 2015). ^{1}H NMR has been applied broadly in wine authentication, due to its ability to provide informative data that could be use in varietal, geographical and vintage authentication (Godelmann et al., 2013; **Table 3**). Gougeon et al. (2018) developed a quantitative NMR (q-NMR) method that can quantify diverse metabolites including alcohols, organic and amino acids, and phenolics, and can be used in combination with chemometrics to differentiate wines based on region and cultivar. However, due to the low sensitivity of NMR, a combination with MS based techniques such as HPLC-MS or GC–MS can be used to provide data on low-level analytes for greater metabolite coverage (Alañón et al., 2015).

Over and above the other techniques, vibrational spectroscopic methods such as IR and Raman are considered as more practicable as they enable in-situ measurements (e.g., with hand held or portable devices) and exhibit sensitivity and robustness

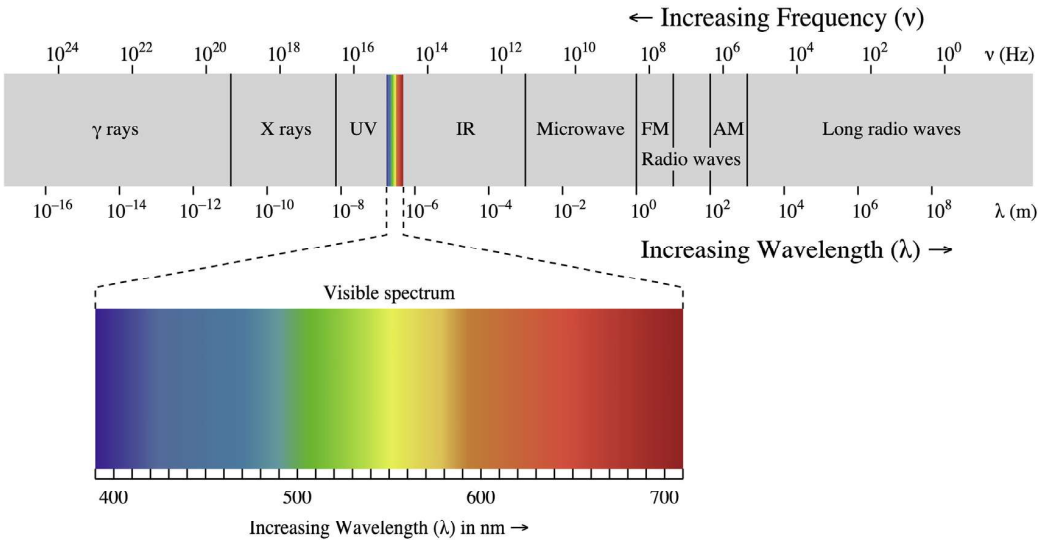


Figure 7 Representation of the electromagnetic spectrum. Reproduced from Ronan P., 2007. EM spectrum.svg, https://upload.wikimedia.org/wikipedia/commons/f/f1/EM_spectrum.svg. Copyright 2007. License at https://creativecommons.org/licenses/by-sa/3.0/.

Table 3 Examples of spectroscopic methods used for wine authentication

Technique	*Chemical marker*	*Classification method*[a]	*Parameters for authentication*	*Reference*
q-NMR	Phenolics, organic and amino acids	PCA, PLS, OPLS	Varietal discrimination of Riesling and Mueller-Thurgau	Ali et al. (2011)
^{1}H NMR	Aromatic region of spectra, organic and amino acids	PCA, LDA, MANOVA	Varietal, geographical and vintage discrimination	Godelmann et al. (2013)
MIR	Carbonyls, organic acids, sugars, alcohols	PCA, LDA	Varietal discrimination	Bevin et al. (2008)
NIR, visible, and UV transmission	Aromatic rings, organic acids, sugars, ethanol	PCA, PLS-DA, SIMCA	Geographical discrimination	Cozzolino et al. (2011)
Vis-NIR	Sugars, ethanol, phenolics, pigments	PCA, PLS-DA	Varietal discrimination	Liu et al. (2006)
Raman fingerprint	FT Raman Sugars, ethanol, phenolics, organic acids	SLDA	Discrimination of wine variety and geographical origin	Magdas et al. (2018b)
	FT Raman Sugars, ethanol, phenolics, organic acids	PCA	Differentiation of grape varieties, production area and aging time	Mandrile et al. (2016)
	SERS Hydroxy-cinnamic acids	SLDA	Discrimination of wine variety and geographical origin	Magdas et al. (2018a)

[a]See Section 3.6 of this chapter for details.

(Teixeira dos Santos et al., 2017a). IR is a versatile technology for routine analysis as it provides the compositional characteristics of a sample by detecting certain functional groups, which can be used for authentication in association with chemometrics (Cozzolino, 2012). The techniques are categorized as near IR (NIR) and mid IR (MIR) according to the range of wavelength. MIR spectroscopy has been employed in distinguishing different varieties (Bevin et al., 2008). Characteristic and specific absorption bands with MIR create more distinct spectra compared to NIR, making MIR more suited to sample identification and quality control (Cozzolino, 2012). Other than wine, MIR has been applied to discriminate Chardonnay juices from different regions (Gambetta et al., 2019). However, calibration and spectral interpretation have been identified as challenges associated with MIR (Cozzolino, 2012). NIR has also been applied in authentication, mostly in combination with visible spectroscopy. A vis-NIR spectroscopic method has been developed to discriminate wines from different geographical regions (Liu et al., 2006; **Table 3**). Moreover, vis-NIR has been implemented in the development of the "BevScan" (**Fig. 8**), which can analyze wine directly through the bottle and has in-built calibrations and classification functions. This type of truly non-destructive technique has the potential to be used in quality control as well as in authentication.

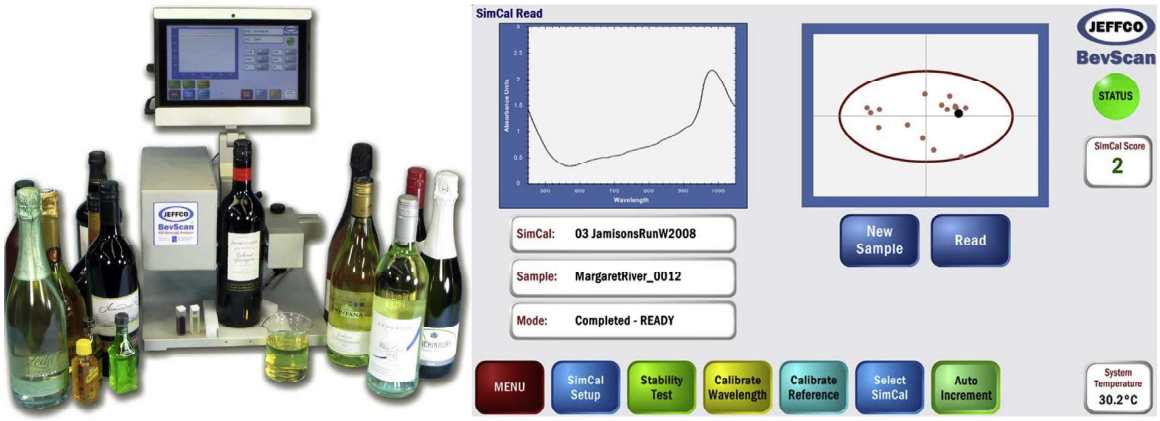


Figure 8 Bevscan: Through bottle beverage analyzer and classifier. Images courtesy of Jeffress Engineering, www.jeffress.com.au.

Raman spectroscopy is not an overly popular analytical method for wine analysis due to the fluorescence effect, but has been adapted for authentication in the pharmaceutical and materials science fields (Das and Agrawal, 2011). The relative inactivity of water molecules in Raman analysis is considered an advantage for wine analysis (Teixeira Dos Santos et al., 2017b). The Raman technique relies on inelastic scattering of monochromatic light from a laser in the visible, near infrared, or near ultraviolet range (Fig. 7). There are two key Raman techniques that are widely applied for food authentication purposes, Fourier transform (FT) Raman spectroscopy and surface-enhanced Raman spectroscopy (SERS) (Craig et al., 2015).

Several studies have evaluated Raman spectroscopy for wine authentication. Mandrile et al. (2016) employed FT Raman to discriminate Italian wine according to variety, vintage and production area (**Table 3**). For white wine discrimination, Magdas et al. (2018a) utilized the SERS technique, determining that it was more successful for the classification of geographical origin rather than vintage. They also explained that according to their previous studies, FT Raman technique provides signals according to the concentration of molecules in a sample whereas SERS gives selective, chemically-specific signals related to particular compounds, which result in different intensities for similar compounds (Magdas et al., 2018a). SERS fingerprinting with minimal sample preparation has been applied to Portuguese white wine, with different incubation times permitting discrimination by variety as well as region (de Almeida et al., 2019).

Chemometrics

Multivariate statistical analysis of chemical data (chemometrics) is important in understanding the relationship between the chemical properties of substances and parameters such as quality or variety. Chemometrics is an important element in authentication, where the identified variations/similarities can be applied when classifying samples according to their origin or variety, or be used to identify non authentic/adulterated samples. Chemometrics can be divided into three categories according to the purpose of analysis: exploratory, classification/discrimination, and regression/prediction (Trygg et al., 2006). Different statistical tools can be applied to classify wines according to their chemical or spectroscopic attributes obtained from different analytical tools, such as discussed in Section "**Analytical Techniques and Sensors**."

Unsupervized methods including principal component analysis (PCA) and cluster analysis (CA) are exploratory, and help to understand the fundamental structure of a dataset. They are used as a pre-processing step for the interpretation of spectra and quantitative data from targeted analyses to discover natural groupings of analytes (Cubero-Leon et al., 2014; **Tables 1–3**). In PCA, the dominant patterns present within samples and variables are illustrated by plotting the scores and loadings matrices. Multiple factor analysis (MFA) is a variation of PCA used for analyzing multiple sets of data, and is important for understanding the impact of different variables, such as elements and volatile compounds, used in a classification (**Fig. 9**). In the same manner, CA allows grouping of samples based on their similarities or differences using all the variance contained in the original data set. Specifically, hierarchical cluster analysis (HCA) has been applied in wine authentication (**Table 1**), to statistically categorize wine into different groups.

Supervised techniques mostly used in authentication include classification models such as linear discriminant analysis (LDA), soft independent modeling of class analogy (SIMCA) and canonical discriminant analysis (CDA) methods (**Tables 1–3**). Classifications are based on preliminary information such as geographical origin or variety, which will serve as a reference for the target results. Among the types of models, discriminant analysis (DA) methods are applied where there are definitive classes. LDA determines the variables with a major discriminant capacity, for instance in geographical authentication by elemental analysis (e.g., ICP-MS, **Table 1**), where stepwise LDA (SLDA) has shown 100% discrimination of red wines from three Australian regions (Martin et al., 2012). CDA is used as a pattern recognition technique that considers the between class variation. For example, CDA has been applied in classifying Austrian wine according to region, variety and vintage using the phenolic pattern (Jaitz et al., 2010; **Table 2**). Classification with SIMCA focuses on similarities within the classes in assigning groups and is widely applied in spectroscopic data analysis (Oliveri and Simonetti, 2016).

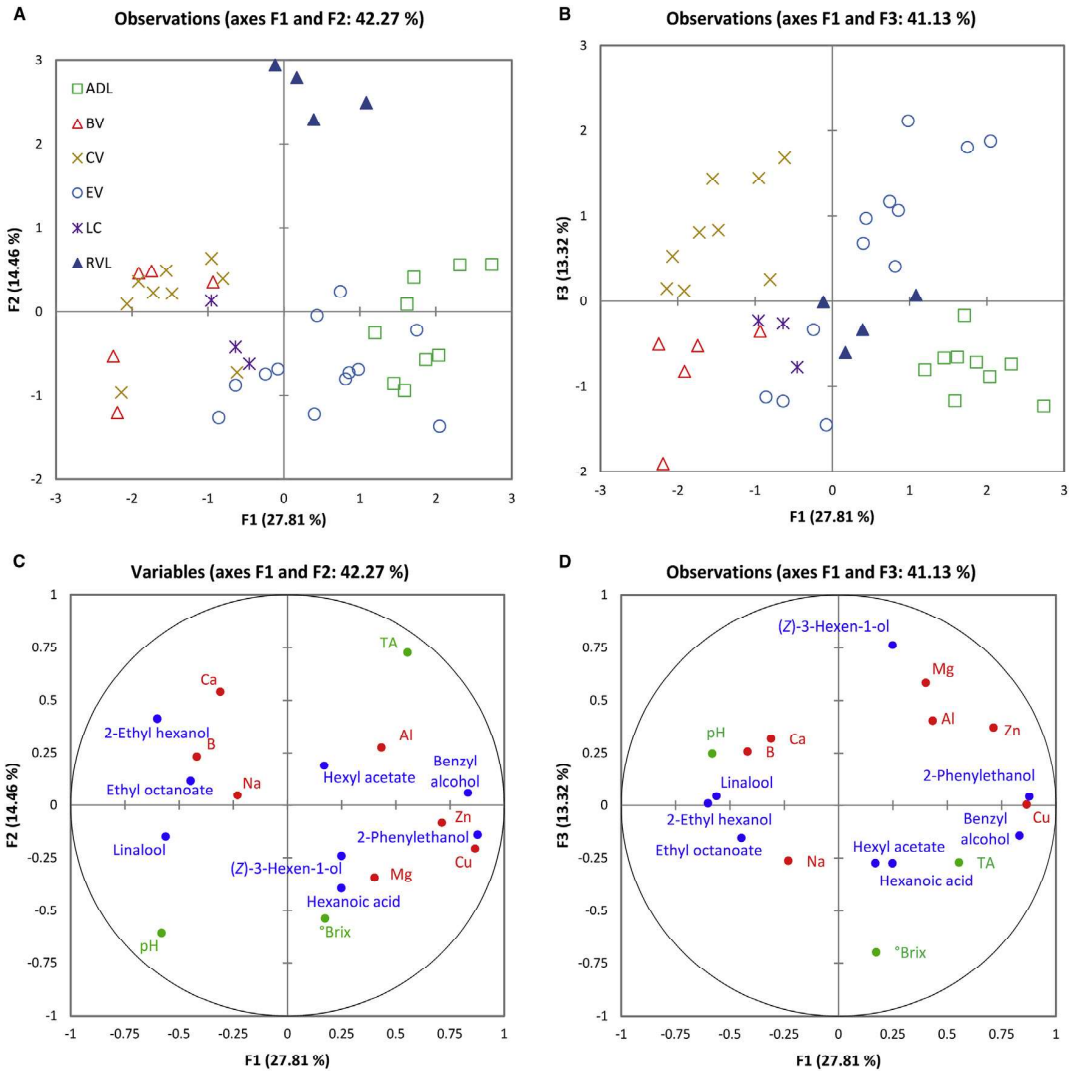


Figure 9 Geographical classification of Chardonnay grape samples from different regions using MFA on various volatiles, basic juice parameters (titratable acidity, pH, total soluble solids), and elements, showing scores on F1 and F2 (A) and F1 and F3 (B), and corresponding loadings (C and D). Reproduced from Gambetta, J.M., Cozzolino, D., Bastian, S.E.P., and Jeffery, D.W., 2017. Exploring the effects of geographical origin on the chemical composition and quality grading of *Vitis vinifera* L. cv. Chardonnay grapes. Molecules, 22:218. Copyright 2017. License at https://creativecommons.org/licenses/by/4.0/.

As a quantitative regression modeling method, partial least squares (PLS) regression is used in differential sorting of many input variables, such as detecting adulterated samples from spectroscopic data (Oliveri and Simonetti, 2016). PLS can be used in combination with discriminant analysis (PLS-DA). The aim is to optimize separation between different groups characterized by several quantitative variables, which can be evaluated by their ability to predict new and unknown samples (Cozzolino, 2012). Classification of Tempranillo wines according to their geographical origin was undertaken, with PLS-DA models showing 100% correct classification of Australian wines and 85% for Spanish samples (Liu et al., 2006). Orthogonal projections to latent structures discriminant analysis (OPLS-DA) is a variation of PLS-DA that constructs more easily interpretable models that are applied in authentication, such as when differentiating samples according to variety by UHPLC-Q-TOF-MS (Rubert et al., 2014; **Table 2**).

For the developed models, validation is conducted by various methods. Analysis of variance (ANOVA) is a data driven validation method that is commonly used in authentication (Westad and Marini, 2015). Cross validation, single evaluation or repeated evaluation are also applied using a training set and a test set, to support the classification of unknown samples (Oliveri and Simonetti, 2016). In selecting validation methods, it is important to consider the aspects of validation such as the robustness, which dictates the ability to predict future samples, or according to overall conclusions such as finding a significant variable (Westad and Marini, 2015).

Wine Traceability

Overview of Traceability

The importance of traceability in the food industry has increased in recent years due to rising consumer awareness of food safety and quality. It is considered as a risk management tool that can be used in the food supply chain as evidence for consumers and producers to verify the origin and overall production process (Vukatana et al., 2016). Traceability of a food, as defined by the European Commission (EC) within Regulation EC No. 178/2002, implies "the ability to trace and follow a food, feed, food-producing animal or substance intended to be, or expected to be, incorporated into a food or feed, through all stages of production, processing and distribution" (European Parliament, 2002). A proper traceability system gives the opportunity to control products as individuals or a batch, and isolate them if issues with food safety or quality arise. It enables the identification of the source of a contamination, for example, which ensures an effective recall procedure. Additionally, a traceability system provides transparency in the supply chain, enabling detection of where items are located (Kok et al., 2012). Therefore, traceability benefits consumers by protecting against fraud and helps optimize and control the production process for producers, such as in a recall situation (Pascu, 2013).

Traceability in a food manufacturing environment can be applied through the production chain from harvest to sale (chain traceability) or within a single step of the chain (internal traceability) (Moe, 1998). Also, according to the objective of the data collection, traceability can be categorized in different ways, mainly as conventional traceability, genetic traceability, or geographic traceability (Dalvit et al., 2007). Application of these methods varies depending on the industry:

- Conventional traceability is primarily applied in processed food systems, using labeling methods;
- Genetic traceability is important in animal production, for verification of breeds through DNA analysis;
- Geographical traceability is aimed at verification of product origin claims, through chemical analysis that yields a profile of elements, volatiles or stable isotopes (Kok et al., 2012).

The latter kind of verification is largely relevant for specialty products with a PDO or Protected Geographical Indication (PGI), such as wines that are linked to regions of origin (Villano et al., 2017).

Application of Wine Traceability

The implementation of a traceability system in the wine industry is critical, given the rapid increase in counterfeiting. Traceability prevents fraud in the supply chain by bringing transparency to the overall process, from raw materials to finished products, and thereafter in distribution and sales (Biswas et al., 2017). Global Standards One (GS1) is a world recognized organization responsible for the development of a global traceability model, and for introducing the Wine Traceability Working Group in 2003 (Cimino and Marcelloni, 2012). The collaboratively developed traceability model emphasizes the responsibilities of each key group in the supply chain (**Table 4**) to enhance the traceability of the system (Wine Traceability, 2018). To ensure tracking through the supply chain, it is important that each key group records all the required information so that the next person in the chain can link those data with the related details of their segment (Palade and Popa, 2014).

Addressing the Issues of a Traceability System

In a food supply chain, it is important to identify any issues that disrupt the effectiveness of traceability. In their review, Dabbene et al. (2014) identified four issues that need to be addressed in order to manage a proficient traceability system:

i Food crisis management is a major concern of a traceability system. In a food recall incident, it is essential to have enough data to track backwards to identify the problem as well as to have required data to track forward to promptly withdraw the affected products. Wynn et al. (2011) introduced a model with a process of collecting and exchanging the required data within the system.
ii Traceability of bulk products is another important concern, including in the wine industry where bulk wine may be stored and shipped. The issue with these products is the inherent possibility of contamination in a batch process, where tanks might contain residues of the preceding batch. Introduction of a threshold limit for batch to batch contamination has been applied in such cases (Comba et al., 2013).
iii Quality and identity preservation (attributes such as geographical origin, varietal origin, or specific production method) are major elements in the traceability system. This is more applicable to specialty products associated with PDO and PGI where product quality is linked to those regions, and luxury products where a minimum level of quality is expected.
iv Fraud prevention is also an important component in a traceability system (Dabbene et al., 2014). Having an effective traceability system allows a company to have more control over what is being produced, and especially who is responsible for each production step, which can tighten the gaps in the supply against fraudulent activities.

It is important to appreciate the significance of a traceability system in the wine industry, as the issues discussed above can have a substantial economic impact if problems arise. A proper traceability system such as the Label Integrity Program introduced by Wine Australia (Wine Australia, 2018) is usually based on different technologies (such as those discussed in Section "**Traceability and Authentication Assurance Systems**") but also needs to be supported with objective analytical evaluations for wine authentication (as discussed in Section "**Authentication Methods for Wine**").

Table 4 The key areas in the wine supply chain determined by the GS1 traceability system

Group	*Responsibilities*	*Information*
Grape grower	Production, harvest, delivery of grapes, sampling analysis	Name and address of the vineyard, plot map reference, size of plot/number of vines, vine variety, chemical content of water, annual treatments and any supplier details with products received and batch numbers
Wine producer	Production, manufacture/blending of wine, sampling analysis	Identification of the wine producer, product identification, records of operations, supplier details of additives with batch numbers, quantities of wine dispatched, shipping container identifications
Bulk distributor	Receipt of wine, sampling analysis, blending, storage, dispatch	Identification of the bulk distributor, quantities of wine received, bulk container identifications, global trade item number, batch and quantity numbers
Transit cellar	Receipt of wine, analysis of bulk wine, storage, dispatch	Identification of the transit cellar, container codes, product identification, the quantity of wine dispatched, batch number of each product dispatched
Filler/Packer	Receipt of wine, analysis of samples, filling, packing, storage, dispatch of finished goods	Shipping container identification, identification of filler/packer and batch number of the dry goods (bottles, caps, corks), lot number of shipped items
Finished goods distributor	Receipt of wine, storage, inventory management, dispatch of finished goods	Lot numbers of inbound pallets and cartons from filler/packer and lot numbers of outbound pallets and cartons to retailer
Retailer	Receipt of finished goods, record of received items, sell to the final consumer	Container number, lot number of the components of the cartons, count of the items contained, packaging date and batch/lot number

Traceability and Authentication Assurance Systems

Different methods of traceability management and authentication assurance are available, which can be divided into five groups (Przyswa, 2014a):

i Tamper-proofing features: aspects that guarantee the product integrity from the moment it was bottled;
ii Overt (visible) features: visual components that are accessible for the consumer;
iii Tracking and tracing technologies: systems used in wine industry logistics to ensure that products are traceable;
iv Covert (coded or hidden) features: visual features that can only be accessible through specialized readers;
v Forensic analysis: investigation of physical, chemical and biological markers to confirm the authenticity of the product (i.e., methods presented in Section "**Authentication Methods for Wine**").

Some of the technologies can be classified in more than one category, as elaborated below.

Tamper-Proofing Technologies

Tamper-proofing devices are the first security check-point for the consumer that do not require any specific prior knowledge about their use nor a particular device to interact with them (Przyswa, 2014b). Tamper-proofing technology aims to provide the consumer with an assurance that the wine has not been altered or replaced after it left the production area (Przyswa, 2014a) and it is a simple visual decision for the consumer: the device needs to be intact and not refurbished, otherwise the product is rejected (Przyswa, 2014b). In its simplest form, tamper-proofing usually encompasses seals and closures that are easily recognized by the consumer when the wine has been tampered with (Przyswa, 2014a), such as the wine in **Fig. 10**, which shows a seal and closure that were damaged when the bottle was opened. Other examples include covers used to conceal the top and sides of a wine bottle to prevent the wine from being opened without removing the plastic, which is damaged irreparably in the process (Bullock III, 1985), or a plastic sleeve can be used around a closure that changes color if any damage has occurred (Model, 2005). A commercial example of a seal that, among other features, uses visual changes to demonstrate tampering is the Bubbletag® from Prooftag (**Fig. 11**). Bubbletag® uses *chaosmetric* technology to generate bubbles in a foil tag to reveal evidence of tampering, with the positioning and the size of the bubbles being created randomly and irreproducibly.

A cover can also be transparent so the consumer can inspect the cork for any damage and view any expected trademark (Taylor, 1989). In addition, holograms can be put to use as a clever tamper-proofing method, whereby the seal does not show any sign of reflective color, but once an attempt has been made to remove the adhesive tag, the holographic layer will become apparent, thus revealing attempts of tampering (Palmasi et al., 2001). A commercial example of this technology is the Variogram® from Prooftag (**Fig. 12**), which includes other information like the Bubbletag®, such as brand logo, quick response (QR) code (see Section "**Quick Response (QR) Codes**") and other information.

Figure 10 Example of simple wine bottle seal and closure (A) that shows evidence of damage when the bottle is opened (B). Image courtesy of David Jeffery.

Figure 11 Example of commercially available Bubbletag® tamper evident seal showing various features including (A) brand logo, (B) chaosmetric tag, (C) tag substrate, (D) quick response (QR) code, and (E) batch identification code. Image courtesy of Prooftag, www.prooftag.net.

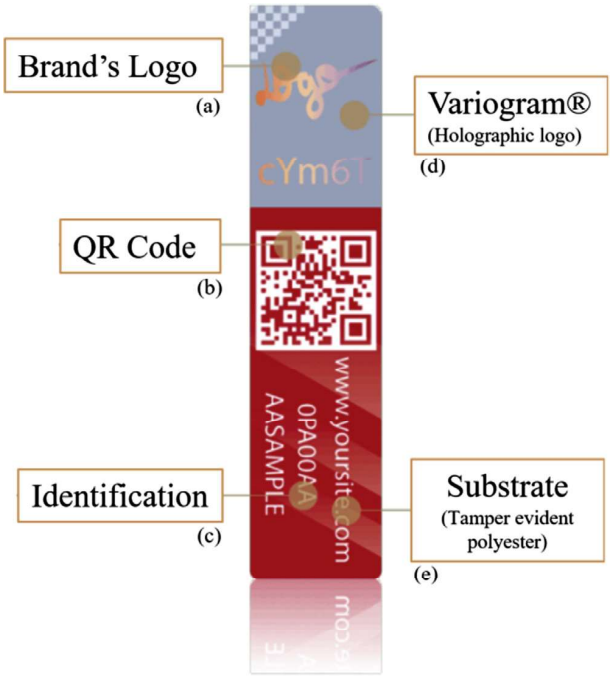


Figure 12 Example of commercially available holographic Variogram® tamper evident seal showing various features including (A) brand logo, (B) quick response (QR) code, (C) batch identification code, (D) holographic tag, and (E) tag substrate. Image courtesy of Prooftag, www.prooftag.net.

Covert Packaging Features

A covert packaging feature is the next step for consumers who are seeking a number of ways to authenticate and trace what they are purchasing (Lecat et al., 2017; Przyswa, 2014a). Covert technology includes visual cues on the label that are difficult to replicate, like holograms, variable inks, films, and watermarks that help the consumer judge the authenticity of the product. **Figs. 13** and **14** show some examples of those features on a spirit bottle that could equally be applied to wine.

A taggant is an example of covert technology that can be used to ensure product authenticity. One variant uses nanoparticles that can be made from a variety of materials designed to be individually encoded and irreproducible (Li, 2013). Microtaggant® (www.microtracesolutions.com) is a commercial example of this technology, which uses articles that can be detected by visual cues with the presence of a specific reader.

Covert technology is a great resource for the consumer to be able to authenticate wine, but as packaging technology for wines constantly evolves, so too does the sophistication of counterfeiting (Lecat et al., 2017). Because of this, it is crucial for the wine industry to keep innovating and relying on a combination of technologies for traceability and authentication (Lecat et al., 2017; Li, 2013). It is therefore important to consider technologies created for other purposes and apply them to the wine industry where appropriate.

Blockchain Systems

Created initially for transactions involving cryptocurrency, blockchain systems utilize a network that functions without requiring trust between the parties involved and without requiring an intermediary body, making traceable transactions safer and faster (Christidis and Devetsikiotis, 2016). Due to these advantages, research into the applicability of blockchain methods for food traceability has been increasing over time, including within the wine industry (Rose, 2019; Biswas et al., 2017) for track and trace technologies used in the supply chain.

A blockchain system was effectively developed for the wine production chain in a study by Biswas et al. (2017). The approach integrated the transactions of all main entities (i.e., grape grower, wine producer, bulk distributor, transit cellar, filler/packer, finished goods distributor, wholesaler and retailer) with each one being represented by a block (Biswas et al., 2017). Every block created two keys, one public and one private, whereby the public key was used to share with all the blocks involved (Biswas et al., 2017). Every bottle of wine was assigned with a unique identification (ID) number, making it impossible to sell the same item more than once and eliminating the possibility of counterfeiting (Biswas et al., 2017). Blockchain also provides traceable information for the consumer who enters the ID number of the purchased bottle into the system, enabling verifiable provenance and evidence of

Figure 13 Examples of commercially available covert and overt technologies (front of bottle) showing (A) tamper-proofing seal, (B) near field communication (NFC) tag, and (C) variable data printing. Images courtesy of All4Labels, www.all4labels.com.

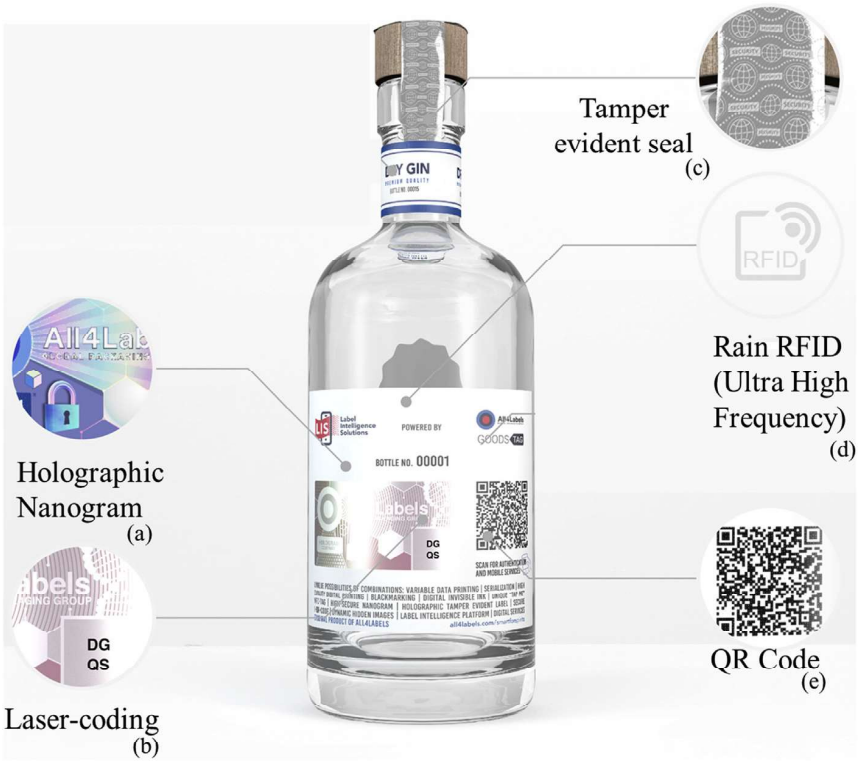


Figure 14 Examples of commercially available covert and overt technologies (rear of bottle) showing (A) holographic label, (B) laser-coding printing, and (C) tamper evident seal, (D) RFID feature, (E) QR code. Images courtesy of All4Labels, www.all4labels.com.

authenticity (Biswas et al., 2017). This provides transparency, safety and security not only for the entities involved in the supply chain but also for the consumer.

Despite the fact that it is a fairly new concept, the advantages of blockchain have garnered interest from companies that are either applying it or are on the path to applying it in their wine business, e.g., Wine Chain (www.winechain.org), Tao Chain (www.dataokeji.cn) and Ezlab (http://www.ezlab.it/case-studies/wine-blockchain).

Radio-Frequency Identification (RFID) Technology

RFID is another technology that has become increasingly popular over the years for facilitating fast handling and management over industrialized products (Want, 2006). RFID is used for in-situ tags that operate on a radio frequency system (mainly ultra-high frequencies), enabling discriminative identification from a distance without human assistance (Want, 2006). Each tag is not only capable of storing the product's ID but can also carry extra information about the product (manufacturer, product type, etc.) and even external conditions (temperature, humidity, etc.) during storage (Want, 2006).

This technology has been studied and integrated into different stages of production and distribution of wine. RFID tags can be placed in the neck of the filled bottle and recognized from every side from a range of around 3–7 m (Xi and Ye, 2012). Tags can also be incorporated into the bottle closure (Hu and Cole, 2010; Gonçalves et al., 2014) and act not only as a covert feature but also as a tamper-proofing mechanism. This technology can be used for winemaking process management (Taylor et al., 2006) and through the entire supply chain, working as a tracking and tracing technology as well (Wang et al., 2012; Fu et al., 2009; Cuiñas et al., 2011; Cimino and Marcelloni, 2012).

Near Field Communication (NFC)

Derived from RFID, NFC is a covert feature that can transmit encoded data based on short range radio communication between two compatible devices in order to conduct a transaction (see **Fig. 13**B; Du, 2013). Yiu (2014) proposed implementing NFC technology as a ready-to-use tool for the consumer to authenticate wine before purchase at a retailer. Using NFC tags on the bottles (**Fig. 13**) and an NFC-enabled smartphone, the consumer can promptly be connected with data collected by the winemaker and be assured that the wine is authentic (Yiu, 2014). The tags can be designed to be breakable and become corrupted once the wine is opened (Yiu, 2014). As with RFID, NFC systems can also be developed to authenticate and record data throughout the wine supply chain (Yiu, 2014). The NFC-enabled anti-counterfeiting system proposed by Yiu (2014) incorporates:

i NFC tags to attach to bottles;
ii NFC compatible smartphones or tablets;
iii mobile application for the end-user to read the NFC tag before purchasing the wine;
iv internal system for the winemaker to manage and store data about the wine;
v mobile application for the winemaker to encode the NFC tag at wine bottling.

Technology that uses radio frequency communication is still evolving but faces technical difficulties when dealing with complex processes, especially involving metals and water (Hu and Cole, 2010). There is also the increased cost of implementing that technology, which could explain why the industry is still relying on simple and low costs systems (Badia-Melis et al., 2015), like barcodes.

Barcodes Systems

Most people would be familiar barcodes that are used on many goods, including on a wine back label, which enable efficient, accurate and automated collection of data (Youssef and Salem, 2007). Barcodes are a covert feature that provide a visual representation of information on a surface that can be interpreted by scanning with a laser bar code reader (at a close distance) (Youssef and Salem, 2007). The first barcodes contained information within parallel printed lines and were called linear barcodes because they could only be read horizontally (e.g., EAN-13, interleaved 2 of 5, and Code 39, **Fig. 15**A–C) (Youssef and Salem, 2007; Noraziah et al., 2011). This has evolved into different formats and patterns that can be read in two dimensions, hence these are known as 2D barcodes (e.g., Maxicode and PDF-417, **Fig. 15**D and E) (Youssef and Salem, 2007; Noraziah et al., 2011). Barcoding can be used as simple and inexpensive tracking and tracing technology (Tzoulis and Andreopoulou, 2013) but it loses its efficiency because of the necessity of having a continuous path between the barcode and the reader (Badia-Melis et al., 2015).

Quick Response (QR) Codes

As another example of a 2D barcode, QR codes provide a larger set of data in less space, mainly because they are interpreted vertically and horizontally as explained above (**Fig. 15**D and E; Brabazon et al., 2014). Created as a solution for supply chain management as a tracking and tracing technology, it has developed into a powerful marketing mechanism (and covert feature) that provides the consumer with instant and detailed information about the wine before purchase (e.g., **Figs. 11**D, **12**B, and **14**E; Higgins et al., 2014). Vukatana et al. (2016) have proposed implementing a QR code system to manage the wine supply chain, and winemakers were deemed to be interested in such an approach because of the low cost of integration. It comes with the added advantage of building trust with consumers as a result of transparency with respect to product traceability.

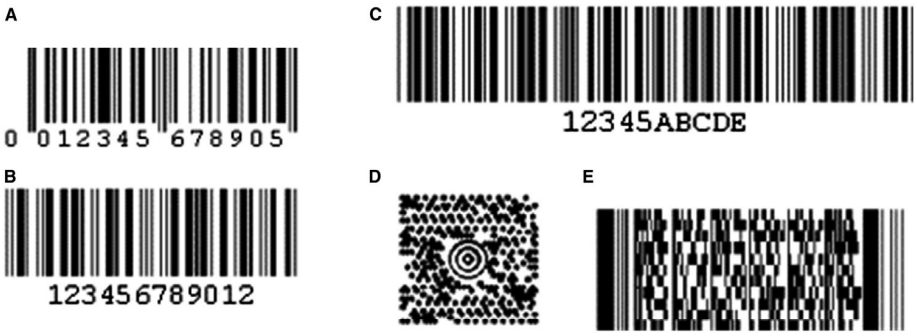


Figure 15 Different types of barcodes: (A) EAN-13, (B) interleaved 2 of 5 (Code 25, I2of5, ITF, I25), (C) Code 39, (D) MaxiCode and (E) PDF-417. Reprinted from Youssef, S.M. and Salem, R.M., 2007. Automated barcode recognition for smart identification and inspection automation. Expert Syst. Appl. 33, 968–977, Copyright (2019), with permission from Elsevier.

Like barcodes, QR codes can be an accessible solution to improve the control in the supply chain and serve as an authentication method for the consumer, but there are limitations. QR codes by themselves do not provide an automatic data feed and require back-end infrastructure to provide the necessary information for their effectiveness. To overcome such limitations, there is research on sensors that can collect and transmit data, and act as a traceability checkpoint.

Wireless Sensor Network (WSN)

Some decisions made within the supply chain might be time sensitive and dependent on real-time local condition measurements, making the development and implementation of wireless sensors an attractive option for controlling the supply chain (Anastasi et al., 2009). WSN is usually composed of a number of interconnected sensor nodes that monitor and report data wirelessly without the need for extensive infrastructure (Yick et al., 2008). The sensors have restricted processing and computing capabilities, and are small and low cost (Yick et al., 2008). They are capable of measuring, sensing, and collecting data about the surroundings, and transferring that data to an operator (Yick et al., 2008).

This technology has been used for vineyard tracking and management (McCulloch et al., 2008; Díaz et al., 2011; Beckwith et al., 2004) as well as managing some stages of wine production, like barrel maturation (Anastasi et al., 2009). In combination with other mature and inexpensive technology technologies like RFID, WSN can be successfully implemented to ensure traceability throughout the wine production supply chain (Expósito et al., 2013). The combination provides continuous monitoring and offers a fast and efficient traceability management system (Expósito et al., 2013). However, one of the challenges faced by Expósito et al. (2013) during implementation was the unfamiliarity of the industry partners with the new system. Training was required, but it was also necessary to adapt the system to better cater to the needs of industry and user support.

In summary, there are a number of technologies available or under development for the industry to be able not only to trace their products but also to offer the consumer transparency and better guarantees of authenticity. However, for the latter to be effective, consumer perspectives of the concepts and technologies need to be taken in account.

Consumer Perspective of Wine Authenticity

Authenticity Concepts and Dimensions

Authenticity can represent different notions when considered from a consumer perspective. There are three authenticity concepts commonly used when studying consumer behavior that can be applied when interpreting wine purchase intentions and attitude:

- Iconic authenticity: covers the idea that a product provides a good image of something else (Grayson and Martinec, 2004) and includes the notion that an object needs to well-represent a pre-existing knowledge of what that object should be (Grayson and Martinec, 2004; Moulard et al., 2015). A product that matches up with a prototype or an exemplar of the category of which it is part, is considered iconic authenticity (Moulard et al., 2015).
- Indexical authenticity: consists of the perception that a product is the original; it is the real one rather than a copy (Grayson and Martinec, 2004). Encompasses the concept that the product is linked to its index, which can be, for instance, a person or a location. It is a concept connected to a physical connotation (Moulard et al., 2015). The uniqueness and rareness of a thing is explained by its association with a specific index (Moulard et al., 2015).
- Moral authenticity: involves the consumer that is less interested in traditions and is looking for a product in which the brand has a genuine intention. These consumers seek companies that do or create something because they deeply believe in it, not because of financial demand (Heine et al., 2016). It transfers the values and meanings embedded in the product (Downing and Parrish, 2019).

Moral authenticity in particular can be a gateway for wineries to engage in an already crowded wine market (Downing and Parrish, 2019). Many of the moral principles already embedded in the winemaking business can be enveloped by moral authenticity, for instance when wineries create stories about environmental sustainable practices around terroir or craftsmanship techniques that convey a sense of moral beliefs with them (Downing and Parrish, 2019).

Authenticity is still an open concept and can be defined differently depending on the product and the target market. Considering authenticity for food products, Lunardo and Guerinet (2007) suggested that authenticity is compiled from three dimensions: projection (self-expression), or in other words, the image that wineries want to project to the consumer and how that image can be associated with their own motives; uniqueness; and originality (respect for tradition and provenance). Wine authenticity can have a strong relationship with quality, provenance, and how far the product is from mass production (Beverland, 2005). For wine authenticity branding, Beverland (2006) suggests six dimensions important to be conveyed to the consumer:

- Heritage and pedigree: heritage is established by connecting the wine brand with its past and by celebrating that history. To demonstrate pedigree, the wineries need to show an ability to consistently track and record their performance by running periodic tastings of past inventories, and by building a reliable stock collection.
- Stylistic consistency: wineries need to deliver a consistent taste in order to give the consumer the ability to form a judgment about the brand independently of the vintage.
- Quality commitments: investing in areas that concern the quality of the wine and conveying that commitment to quality is a critical factor for the consumer.
- Relationship to place: relying on the place of origin is a core strategy for the wine industry and it is no less important when portraying an authentic brand. For the consumer, the place of origin conveys a commitment to terroir, traceable origin and unique product.
- Method of production: the consumer has a need to know what goes into the production of the wine and wineries serve that need by providing the bond between the production process and the final product.
- Downplaying commercial motives: consumers value wines that are less mass-marketed. To have a high valued brand, wineries need to express their ability to always prioritize production of a better quality product over commercial motives.

Even though it is still an open concept, authentic products have increasingly become an important factor when discussing wine (Rössel et al., 2016). For example, analyzing wine journalism in Germany from 1947 to 2008, Rössel et al. (2016) observed that a shift occurred with the language when writing about wine, giving more emphasis to the authentic aesthetics of the wine instead of the technical and commercial angles. Authentic aesthetics of wine can be emphasized through geographical linkages, techniques that encompass craft production, and enhancement of an identity coupled with the winemaker as an individual, along with contextualization of the production as a historical and traditional marker (Rössel et al., 2016).

Looking for Authentic Products: The Consumer Profile

The search for authentic experiences can be considered as a refuge for consumers that invokes an emotional sense of nostalgia (Lunardo and Guerinet, 2007). Production methods that are more traditional and nostalgic are linked to authentic foods, in contrast to foods that are produced through industrialized methods and enhanced agriculture (Sidali and Hemmerling, 2014).

Consumers today display a great awareness of what they consume and the risks associated with food products (Lunardo and Guerinet, 2007). This behavior is led by two principles:

- Classification, in which food is classified so that norms and rules can be created, meaning some foods can be authentic and others not;
- Incorporation, whereby regulating what is eaten provides a level of control that influences self-esteem, and where authenticity may afford a guarantee about what is being consumed (Lunardo and Guerinet, 2007).

In essence, consumers are looking for brands that have the same ideals that they live by (Larson, 2012).

Generally, when considering the endorsement of authentic products, consumers seeking authenticity have a tendency to trust other consumers' judgments more than professional recommendations or commercial marketing strategies (Liao and Ma, 2009). On the other hand, advertisement claims can be used to emphasize an authentic image for the consumer (Beverland et al., 2008). Advertising can transmit three forms of authenticity: pure authenticity (defined in Section "**Authenticity Concepts and Dimensions**" as indexical authenticity), approximate authenticity (defined in Section "**Authenticity Concepts and Dimensions**" as iconic authenticity), and moral authenticity (Section "**Authenticity Concepts and Dimensions**"), and one of those forms will be used primarily to judge a brand's authenticity (Beverland et al., 2008).

Branding Authenticity: What Do Consumers Want From an Authentic Brand?

A fundamental factor for creating a successful brand is authenticity, as it builds a distinctive trademark identity (Beverland, 2005; Kapferer, 2008). In general, when a brand is perceived to be managed by individuals that are truly interested, passionate, and motivated by what they are producing, that brand can be perceived as more authentic (Moulard et al., 2016). The opposite is also true, whereby a brand can be perceived as non-authentic when it is directed by professionals that are driven by profit and commercial

values (Moulard et al., 2016). In conjunction with the wine professional's ideals, their nationality and heritage may also influence the consumer authenticity perception (Spielmann and Babin, 2011). For example, authenticity perceptions of American consumers were highest when the winemaker and the wine had the same heritage, but that statement was only true for wines coming from France in comparison to wines coming from USA (Spielmann and Babin, 2011). Because American consumers have the image that French wines are an exemplary representation of the category, when the appearance of "Frenchness" is reinforced by having a French winemaker producing American wines, the value of those wines was also enhanced (Spielmann and Babin, 2011).

Presenting an image of an authentic brand to the consumer is crucial for wineries to strengthen their position as a quality brand, to mandate premium prices, and to differentiate themselves from the competition (Beverland, 2005; Sjostrom et al., 2014). In order to offer an authentic image, wineries need to establish and promote a truthful and sincere brand story. This is accomplished by considering the six dimensions defined in Section "**Authenticity Concepts and Dimensions**"; in order words, by linking the wine to its provenance, reinforcing traditional methods, demonstrating a passion and commitment for the craft, having quality production rather than mass industrialized production, paying attention to detail, and always appearing to put ideals above commercial motives (Beverland, 2005, 2006). Increasing the authenticity perception of a wine and the associated brand not only provides differentiation between brands, especially for craft industries (Downing and Parrish, 2019), but also provides producers with the opportunity to increase the perceived value of their wine (Moulard et al., 2015).

Influence of a Wine Label on Authenticity Perception

As outlined in Section "**Wine Quality**", wine quality is a complex concept that can be perceived through intrinsic and extrinsic cues. Extrinsic cues are of great significance for consumers because purchase decisions are usually made without the opportunity of prior tasting (Sáenz-Navajas et al., 2013). Important extrinsic cues in relation to perceived quality include place of origin, regulatory denominations around origin (appellations), label aesthetic, and bottling process (on- or off-site bottling) (Sáenz-Navajas et al., 2013). These cues afford a wine label a very prominent role in consumer purchase decision making (Thomas and Pickering, 2003) and have a close relationship with the authenticity perceived by consumers (Lunardo and Guerinet, 2007). Indeed, for younger consumers (between 18 and 25 years) in particular, a label can be the decisive signal for authentic versus non-authentic wines, due mainly to the connection of the wine with where it comes from and where it was produced (i.e., indexical authenticity) (Lunardo and Guerinet, 2007).

The description of terroir on a label can also have an impact on authenticity perceptions, depending on the country of origin (Moulard et al., 2015). New World wines were perceived as more authentic when the labels displayed a highly specific description about the terroir, whereas for Old World wines, a vague terroir description was linked with increased perceptions of authenticity (Moulard et al., 2015). The vague description was perceived as authentic because consumers evaluating Old World wines are easily connected to those countries that are highly emotionalized and a common tourist destination with deep cultural linkages. It is likely that a pragmatic description about terroir for Old World wine weakens the transcendent experience (Moulard et al., 2015). Interestingly though, label innovation for Old World wines can be challenging, as in the case of consumers comparing two French wine labels (one modern and one traditional), where the modern label was perceived as a risk and not authentic (Lunardo and Guerinet, 2007).

For Old World wine countries, the geographical origin of the wine is a major factor for authenticity and the connection between the wine and its production origin is regulated through a legislative scheme called PDO (Section "**Applications of Wine Authentication**"; Kamiloglu, 2019). Indeed, the use of the logos indicating the PDO on the label of a wine was positively correlated to the concepts of a protected/delimited region, terroir, and authenticity (Ginon et al., 2014). Terroir can also add an intangible concept of nature giving a "perception of space" (Capitello et al., 2016). It conveys to the consumer cues to assist in building an image of uniqueness, identity, and authenticity (Charters, 2010; Beverland, 2006). When terroir is used as a marketing tool targeting the consumer, it intends to offer a symbolic connotation of authenticity, and is used to segregate the sense of industrial wine from "genuine" wine (Charters, 2010) and as a guarantee of an authentic offering (Gade, 2004).

Wine Authenticity and Innovation

For a product to stay current in a competitive market and attend to a consumer's desires and expectations, product innovation is crucial for any business (Armstrong and Kotler, 2012). However, because the wine industry is usually tightly connected to traditional production methods and a strong concept of terroir, innovation can have an impact on the consumer's perceptions of authenticity (Spielmann and Charters, 2013; Qesja et al., 2016). Indeed, even though there is a shift in wine trade in terms of the dominance of Old World versus New World wines as discussed in Section "**Wine as a Global Beverage**", consumers are willing to pay more for Old World wines because of the perception that they are more authentic (Moulard et al., 2015). Old World wines apparently match the expectations that consumers have of what the "true" wine represents, according to Moulard et al. (2015). This might be a consequence of New World wine producing countries being more open to adopting innovative strategies, whether in the vineyard or with winemaking practices (Aylward, 2003).

When it comes to implementing new technology in the wine industry, for example, nanotechnology, consumers were observed to be cautious with the application of such technology (Chiodo et al., 2015). When comparing a wine labeled as "nano wine" with another with no designation, the "nano wine" had an overall consumption rejection (Chiodo et al., 2015). However, with a deeper comprehension about the application of nanotechnology, consumers were more open to the concept when it was used to enhance

the authenticity of the wine, such as using nanotechnology devices (like taggants, Section "**Covert Packaging Features**") to improve surveillance systems and tracking movements through the supply chain (Weiss et al., 2006). QR codes and NFC tags are another example of technologies that are increasingly presented to the consumer on wine packaging (Sections "**Quick Response (QR) Codes**" and "**Near Field Communication (NFC)**"). For consumers interested in information about regionality and environmental-friendly processes involved in the production of the wine, having a QR code present on the label gives a positive impression (Higgins et al., 2014) and also increases the expected price of the wine (Sillani et al., 2017). In the case of NFC tags, whether consumers intend to use it to authenticate wine will mainly depend on two determinants: the technology is safe from privacy problems, and it fits their lifestyle and fulfills their requirement for information (Bandinelli et al., 2017).

Another factor that needs consideration is the level of involvement of the consumer, either within a country or as an individual. Qesja, Crouch and Quester (2016) observed that for lower alcohol wines, which can be produced using various winemaking technologies, in a country such as Indonesia where wine is not considered a traditional product by the majority of the participants, an innovative wine (with no alcohol content in this case) was positively reacted to and was still considered authentic. When the same study was conducted in Australia, where wine is considered as a traditional product by the majority of the participants, the innovation was perceived negatively and was not considered to be authentic (Qesja et al., 2016).

Consumer Perspective on Traceability

As discussed in Section "**Wine Traceability**", traceability processes and devices provide consumers with a means to connect the final product with the raw material and with the producer, helping to both guarantee and satisfy the need for an authentic product (Dimara and Skuras, 2003). The concept of traceability in food products can still be confusing, although generally, when consumers know what traceability is, they connect it to identifying the origin of the product and the production process. Nonetheless, this can be different across countries and across different products (Kehagia et al., 2007).

When traceability is used as an extrinsic quality cue to influence consumer wine purchase decisions, it is usually to target consumers seeking information on a wine label about tradition and authenticity (Dimara and Skuras, 2003). Indeed, according to Kehagia et al. (2007), the best way to communicate traceability to the consumer is through the product's labels and quality seals. However, these need to be simple and clear to facilitate understanding, whereas more complex communication methods that require specific knowledge and reading devices for barcodes or electronic tags might only be confusing and decrease consumer confidence in the product.

Conclusion and Future Perspectives

Wine has long been a part of human history and today it represents a culture marker in society. It is compositionally complex and highly connected with its place of origin and with different quality parameters, which all play an important role in its perceived value. Wine is also a luxury product in a global marketplace, and is perhaps more subject to fraud and adulteration than other foods. This is a concern for the consumer and has great financial implications for the wine producers, making wine authentication and traceability important tools.

It is a part of any quality management system to ensure the safety and authenticity of the products being produced. To achieve that, wine authentication needs to be addressed with a multi-dimensional perspective, using different evaluation methods. The development of different analytical approaches that provide a measure of authenticity, when paired with traceability systems and devices, can be used to investigate and guarantee a safe and unadulterated wine for the consumers.

Analytical approaches are based on identification and quantification of specific chemical markers. Using omic approaches, targeted and non-targeted methods can be applied in the process of authentication from the level of a gene to that of molecules, but mostly falling within metabolomics and isotopolomics, and utilizing mass spectrometry in conjunction with chromatographic instruments. Fingerprinting methods based on mass spectrometry and spectroscopic techniques are promising in terms of the speed of analysis, and for detecting certain chemical markers for authentication and identifying adulteration.

For further methodological development, a systematic approach is required that considers chemical compositional aspects and takes into account the applicability of these techniques throughout the supply chain. In that sense, the most appropriate option would be inexpensive methods such as sensor techniques, which give non-destructive, rapid results, and user-friendly hand held instruments that could be easily used in the field. This would help in overcoming limitations in the audit process in the wine industry, as currently only a fraction of commercial wines are chemically audited. Development of more sensitive methods with a higher selectivity would be valuable; for example, in geographical origin authentication, the effect of close proximity of regions has limited the efficiency of the current techniques such as stable isotope analysis by IRMS and elemental analyses by ICP–MS.

With all of these techniques, the application of chemometrics plays a crucial role, since it allows the extraction of hidden information from the acquired multidimensional data, as well as the creation of prediction models in order to classify unknown samples. Commonly applied unsupervized methods (PCA, CA) and supervised methods (LDA, PLS-DA, SIMCA) are currently used for discrimination, classification, modeling and correlation. Validation to ensure robustness of models is an important aspect as well.

Having an appropriate traceability system in place is another way to assure wine authenticity. This system essentially involves record keeping throughout the supply chain, which can be effected by using technological devices. Tags, seals and specific markers

on bottle labels can be used not only to ensure effective supply chain tracking but also to guarantee the consumer confidence when the purchase decision is being made. For consumers that are eager for more information about the product or want to verify its provenance, features like NFC and QR codes embedded on the label are an efficient solution. Additionally, those features can be used within the wine supply chain to track the product and to keep records throughout production.

As consumers are the ultimate stakeholders, it is important for producers to consider that authenticity can have different meanings for the consumer and can be used as a marketing strategy by the wine industry. Wine can convey authenticity by creating and transmitting a sincere story that connects the wine with its provenance, with traditional production practices, and with the moral beliefs of the producer. That can be accomplished through information offered to the consumer on the label or through brand construction. Additionally, for effectively transmitting an image of authenticity, the wine industry has to always keep in mind their target consumer, their reputation as a wine producing country, and innovations they are willing to implement.

Acknowledgments

R.K.R.R. acknowledges the support received through the provision of an Australian Government Research Training Program Scholarship and a Wine Australia supplementary scholarship (WA Ph1909). The Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au; project number IC170100008) is funded by the Australian Government with additional support from Wine Australia, Waite Research Institute, and industry partners. L.S.G. acknowledges Wine Australia for providing a supplementary scholarship (WA Ph1804).

References

Alañón, M.E., Pérez-Coello, M.S., Marina, M.L., 2015. Wine science in the metabolomics era. Trends Anal. Chem. 74, 1–20.

Ali, K., Maltese, F., Toepfer, R., Choi, Y.H., Verpoorte, R., 2011. Metabolic characterization of palatinate German white wines according to sensory attributes, varieties, and vintages using NMR spectroscopy and multivariate data analyses. J. Biomol. NMR 49, 255–266.

Anastasi, G., Farruggia, O., Lo Re, G., Ortolani, M., 2009. Monitoring high-quality wine production using wireless sensor networks. In: Sprague Jr., R.H. (Ed.), Proceedings of 2009 42nd Hawaii International Conference on System Sciences. Institute of Electrical and Electronic Engineers, Big Island, USA, pp. 298–305.

Anderson, K., Nelgen, S., Pinilla, V., 2017. Wine markets by country: decadal data, 1860s to 2000s, and 2010-16. In: Global Wine Markets, 1860 to 2016: A Statistical Compendium. University of Adelaide Press, Adelaide, Australia, pp. 157–284.

Armstrong, G., Kotler, P., 2012. New product development strategy. In: Principles of Marketing, 5 ed. Pearson Education, UK, pp. 254–267.

Arvanitoyannis, I.S., Katsota, M.N., Psarra, E.P., Soufleros, E.H., Kallithraka, S., 1999. Application of quality control methods for assessing wine authenticity: use of multivariate analysis (chemometrics). Trends Food Sci. Technol. 10, 321–336.

Aung, M.M., Chang, Y.S., 2014. Traceability in a food supply chain: safety and quality perspectives. Food Contr. 39, 172–184.

Aylward, D.K., 2003. A documentary of innovation support among new world wine industries. J. Wine Res. 14, 31–43.

Badia-Melis, R., Mishra, P., Ruiz-García, L., 2015. Food traceability: new trends and recent advances. A review. Food Contr. 57, 393–401.

Bandinelli, R., Fani, V., Rinaldi, R., 2017. Customer acceptance of NFC technology: an exploratory study in the wine industry. Int. J. Real. Ther. 8, 1–16.

Banks, G., Overton, J., 2010. Old World, New World, Third World? Reconceptualising the worlds of wine. J. Wine Res. 21, 57–75.

Barrias, S., Fernandes, J.R., Eiras-Dias, J.E., Brazão, J., Martins-Lopes, P., 2019. Label free DNA-based optical biosensor as a potential system for wine authenticity. Food Chem. 270, 299–304.

Beckwith, R., Teibel, D., Bowen, P., 2004. Unwired wine: sensor networks in vineyards. In: Proceedings of IEEE Sensors 2004, vol. 2, pp. 561–564.

Bentlin, F.R.S., Pulgati, F.H., Dressler, V.L., Pozebon, D., 2011. Elemental analysis of wines from South America and their classification according to country. J. Braz. Chem. Soc. 22, 327–336.

Berna, A.Z., Trowell, S., Clifford, D., Cynkar, W., Cozzolino, D., 2009. Geographical origin of Sauvignon Blanc wines predicted by mass spectrometry and metal oxide based electronic nose. Anal. Chim. Acta 648, 146–152.

Beverland, M., 2006. The 'real thing': branding authenticity in the luxury wine trade. J. Bus. Res. 59, 251–258.

Beverland, M.B., 2005. Crafting brand authenticity: the case of luxury wines. J. Manag. Stud. 42, 1003–1029.

Beverland, M.B., Lindgreen, A., Vink, M.W., 2008. Projecting authenticity through advertising: consumer judgments of advertisers' claims. J. Advert. 37, 5–15.

Bevin, C.J., Dambergs, R.G., Fergusson, A.J., Cozzolino, D., 2008. Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. Anal. Chim. Acta 621, 19–23.

Biswas, K., Muthukkumarasamy, V., Tan, W.L., 2017. Blockchain based wine supply chain traceability system. In: Proceedings of the 2017 Future Technologies Conference (FTC). Vancouver, Canada The Science and Information (SAI) Organization, pp. 56–62.

Borneman, A.H., Herderich, M., Johnson, D., 2013. The DNA of innovation. Wine Vitic. J. 28, 52–56.

Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., Busto, O., 2015. Data fusion methodologies for food and beverage authentication and quality assessment – a review. Anal. Chim. Acta 891, 1–14.

Brabazon, T., Winter, M., Gandy, B., 2014. Digital Wine: How QR Codes Facilitate New Markets for Small Wine Industries. Springer, Singapore.

Bullock III, J.J., 1985. Wine Bottle Cover, 4,506,797.

Campbell, G., Guibert, N., 2006. Introduction: Old World strategies against New World competition in a globalising wine industry. Br. Food J. 108, 233–242.

Capitello, R., Agnoli, L., Charters, S., Begalli, D., 2016. Terroir in a bottle: segmenting consumer choices in generation Y. In: Petruzzellis, L.,W.R. (Ed.), Proceedings of the Rediscovering the Essentiality of Marketing. Springer International Publishing, Cham, Switzerland, pp. 517–528.

Carcea, M., Brereton, P., Hsu, R., Kelly, S., Marmiroli, N., Melini, F., Soukoulis, C., Wenping, D., 2009. Food authenticity assessment: ensuring compliance with food legislation and traceability requirements. Qual. Assur. Saf. Crop Foods 1, 93–100.

Cavicchi, A., Santini, C., 2011. Brunellopoli: a wine scandal under the Tuscan sun. Tourism Rev. Int. 15, 253–267.

Cereda, A., Kravchuk, A.V., D'Amato, A., Bachi, A., Righetti, P.G., 2010. Proteomics of wine additives: mining for the invisible via combinatorial peptide ligand libraries. J. Proteom. 73, 1732–1739.

Cevallos-Cevallos, J.M., Reyes-De-Corcuera, J.I., Etxeberria, E., Danyluk, M.D., Rodrick, G.E., 2009. Metabolomic analysis in food science: a review. Trends Food Sci. Technol. 20, 557–566.

Charters, S., 2010. Marketing terroir: a conceptual approach. In: Proceedings of The 5th International Academy of Wine Business Research Conference. University of Auckland, Auckland, New Zealand, pp. 8–10.

Charters, S., Pettigrew, S., 2003. The intrinsic dimensions of wine quality: an exploratory investigation. In: Larry Lockshin, C.R. (Ed.), Proceedings of the International Colloquium in Wine Marketing. Wine Marketing Research Group, Adelaide, Australia, pp. 1–21.
Charters, S., Pettigrew, S., 2007. The dimensions of wine quality. Food Qual. Prefer. 18, 997–1007.
Chiodo, E., Casolani, N., Greehy, G., Fantini, A., McCarthy, M., 2015. Consumer perceptions of nanotechnology applications in Italian wine. Ital. J. Food Sci. 27, 221–235.
Christidis, K., Devetsikiotis, M., 2016. Blockchains and smart contracts for the internet of things. IEEE Access 4, 2292–2303.
Christoph, N., Hermann, A., Wachter, H., 2015. 25 years authentication of wine with stable isotope analysis in the European Union - review and outlook. In: Jean Marie, A. (Ed.), 38th World Congress of Vine and Wine, 02020.
Cimino, M.G.C.A., Marcelloni, F., 2012. Enabling traceability in the wine supply chain. In: Anastasi, G., Bellini, E., Di Nitto, E., Ghezzi, C., Tanca, L., Zimeo, E. (Eds.), Methodologies and Technologies for Networked Enterprises: Artdeco: Adaptive Infrastructures for Decentralised Organisations. Springer, Berlin, Heidelberg, Germany, pp. 397–412.
Coetzee, P.P., van Jaarsveld, F.P., Vanhaecke, F., 2014. Intraregional classification of wine via ICP-MS elemental fingerprinting. Food Chem. 164, 485–492.
Comba, L., Belforte, G., Dabbene, F., Gay, P., 2013. Methods for traceability in food production processes involving bulk products. Biosyst. Eng. 116, 51–63.
Cozzolino, D., 2012. Recent trends on the use of infrared spectroscopy to trace and authenticate natural and agricultural food products. Appl. Spectrosc. Rev. 47, 518–530.
Cozzolino, D., Cynkar, W., Shah, N., Smith, P.A., 2011. Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? Food Chem. 126, 673–678.
Cozzolino, D., Smyth, H.E., Cynkar, W., Dambergs, R.G., Gishen, M., 2005. Usefulness of chemometrics and mass spectrometry-based electronic nose to classify Australian white wines by their varietal origin. Talanta 68, 382–387.
Craig, A.P., Franca, A.S., Irudayaraj, J., 2015. Vibrational spectroscopy for food quality and safety screening. In: High Throughput Screening for Food Safety Assessment. Woodhead Publishing, pp. 165–194.
Cubero-Leon, E., Peñalver, R., Maquet, A., 2014. Review on metabolomics for food authentication. Food Res. Int. 60, 95–107.
Cuiñas, I., Catarinucci, L., Trebar, M., 2011. RFID from farm to fork: traceability along the complete food chain. In: Proceedings of the Progress in Electromagnetics Research Symposium. The Electromagnetics Academy, Cambridge, MA, USA, pp. 1370–1374.
Cynkar, W., Dambergs, R., Smith, P., Cozzolino, D., 2010. Classification of Tempranillo wines according to geographic origin: combination of mass spectrometry based electronic nose and chemometrics. Anal. Chim. Acta 660, 227–231.
D'Alessandro, S., Pecotich, A., 2013. Evaluation of wine by expert and novice consumers in the presence of variations in quality, brand and country of origin cues. Food Qual. Prefer. 28, 287–303.
Dabbene, F., Gay, P., Tortia, C., 2014. Traceability issues in food supply chain management: a review. Biosyst. Eng. 120, 65–80.
Dalvit, C., De Marchi, M., Cassandro, M., 2007. Genetic traceability of livestock products: a review. Meat Sci. 77, 437–449.
Danezis, G.P., Tsagkaris, A.S., Camin, F., Brusic, V., Georgiou, C.A., 2016. Food authentication: techniques, trends & emerging approaches. Trends Anal. Chem. 85, 123–132.
Das, R.S., Agrawal, Y.K., 2011. Raman spectroscopy: recent advancements, techniques and applications. Vib. Spectrosc. 57, 163–176.
de Almeida, M.P., Leopold, N., Franco, R., Pereira, E., 2019. Expedite SERS fingerprinting of Portuguese white wines using plasmonic silver nanostars. Front. Chem. 7, 368.
Della Corte, A., Mattivi, F., Vrhovsek, U., 2013. Development of a targeted LC-MS/MS method for the determination of different classes of grape lipids. In: Jeandet, P. (Ed.), Eight Symposium in Vino Analytica Scientia: Analytical Chemistry for Wine, Brandy and Spirits. University of Reims, Reims, France, p. 47 (CA3).
Di Paola-Naranjo, R.D., Baroni, M.V., Podio, N.S., Rubinstein, H.R., Fabani, M.P., Badini, R.G., Inga, M., Ostera, H.A., Cagnoni, M., Gallegos, E., Gautier, E., Peral-García, P., Hoogewerff, J., Wunderlin, D.A., 2011. Fingerprints for main varieties of Argentinean wines: terroir differentiation by inorganic, organic, and stable isotopic analyses coupled to chemometrics. J. Agric. Food Chem. 59, 7854–7865.
Di Stefano, V., Avellone, G., Bongiorno, D., Cunsolo, V., Muccilli, V., Sforza, S., Dossena, A., Drahos, L., Vékey, K., 2012. Applications of liquid chromatography–mass spectrometry for food analysis. J. Chromatogr. A 1259, 74–85.
Díaz, S.E., Pérez, J.C., Mateos, A.C., Marinescu, M.-C., Guerra, B.B., 2011. A novel methodology for the monitoring of the agricultural production process based on wireless sensor networks. Comput. Electron. Agric. 76, 252–265.
Dimara, E., Skuras, D., 2003. Consumer evaluations of product certification, geographic association and traceability in Greece. Eur. J. Market. 37, 690–705.
Downing, J.A., Parrish, D.C.S.C., 2019. Welcome to my house, do you like the neighborhood? Authenticity differentiation within strategic groups of wineries. In: Sigala, M., Robinson, R.N.S. (Eds.), Management and Marketing of Wine Tourism Business: Theory, Practice, and Cases. Palgrave Macmillan, New York, USA, pp. 277–294.
Dreisewerd, K., 2014. Recent methodological advances in MALDI mass spectrometry. Anal. Bioanal. Chem. 406, 2261–2278.
Du, H., 2013. NFC technology: today and tomorrow. Int. J. Fut. Comput. Commun. 2, 351.
Easingwood, C., Lockshin, L., Spawton, A., 2011. The drivers of wine regionality. J. Wine Res. 22, 19–33.
European Parliament, 2002. Regulation (EC) No 178/2002 of the European Parliament and of the Council of 28 January 2002 Laying Down the General Principles and Requirements of Food Law, Establishing the European Food Safety Authority and Laying Down Procedures in Matters of Food Safety.
Expósito, I., Gay-Fernández, J.A., Cuiñas, I., 2013. A complete traceability system for a wine supply chain using Radio-Frequency Identification and wireless sensor networks [Wireless Corner]. IEEE Antenn. Propag. Mag. 55, 255–267.
Fandl, K.J., 2018. Regulatory policy and innovation in the wine industry: a comparative analysis of old and new world wine regulations. Am. Univ. Int. Law Rev. 34, 279–350.
FAO/IAEA, 2017. Food Safety and Traceability [Online]. Food Safety and Traceability. Available: http://www-naweb.iaea.org/nafa/fep/topic-food-safety.html. (Accessed 10 April 2018).
Fu, I.-j., Lo, Y.-C., Chen, Y.-J., 2009. Anti-fake Packaging Material of RFID and its Packaging Method. United States patent application US 2009/0051540 A1. 21/06/2008.
Gade, D.W., 2004. Tradition, territory, and terroir in French viniculture: cassis, France, and appellation contrôlée. Ann. Assoc. Am. Geogr. 94, 848–867.
Gambetta, J.M., Cozzolino, D., Bastian, S.E.P., Jeffery, D.W., 2019. Classification of Chardonnay grapes according to geographical indication and quality grade using attenuated total reflectance mid-infrared spectroscopy. Food Analyt. Methods 12, 239–245.
Garrido, J., Borges, F., 2013. Wine and grape polyphenols — a chemical perspective. Food Res. Int. 54, 1844–1858.
Geana, E.I., Popescu, R., Costinel, D., Dinca, O.R., Ionete, R.E., Stefanescu, I., Artem, V., Bala, C., 2016a. Classification of red wines using suitable markers coupled with multivariate statistic analysis. Food Chem. 192, 1015–1024.
Geana, E.I., Popescu, R., Costinel, D., Dinca, O.R., Stefanescu, I., Ionete, R.E., Bala, C., 2016b. Verifying the red wines adulteration through isotopic and chromatographic investigations coupled with multivariate statistic interpretation of the data. Food Contr. 62, 1–9.
Ginon, E., Ares, G., Laboissière, L.H.E.S., Brouard, J., Issanchou, S., Deliza, R., 2014. Logos indicating environmental sustainability in wine production: an exploratory study on how do Burgundy wine consumers perceive them. Food Res. Int. 62, 837–845.
Gladstones, J., 2011. Wine, Terroir and Climate Change. Wakefield Press, Mile End, South Australia.
Godelmann, R., Fang, F., Humpfer, E., Schütz, B., Bansbach, M., Schäfer, H., Spraul, M., 2013. Targeted and nontargeted wine analysis by 1H NMR spectroscopy combined with multivariate statistical analysis. Differentiation of important parameters: grape variety, geographical origin, year of vintage. J. Agric. Food Chem. 61, 5610–5619.
Gonçalves, R., Magueta, R., Pinho, P., Carvalho, N.B., 2014. RFID passive tag antenna for cork bottle stopper. In: Proceedings of the IEEE Antennas and Propagation Society International Symposium. Institute of Electrical and Electronic Engineers, USA, pp. 1518–1519.
Gougeon, L., Da Costa, G., Le Mao, I., Ma, W., Teissedre, P.-L., Guyon, F., Richard, T., 2018. Wine analysis and authenticity using 1H-NMR metabolomics data: application to Chinese wines. Food Analyt. Methods 11, 3425–3434.
Grayson, K., Martinec, R., 2004. Consumer perceptions of iconicity and indexicality and their influence on assessments of authentic market offerings. J. Consum. Res. 31, 296–312.

Guadarrama, A., Fernández, J.A., Íñiguez, M., Souto, J., de Saja, J.A., 2001. Discrimination of wine aroma using an array of conducting polymer sensors in conjunction with solid-phase micro-extraction (SPME) technique. Sensor. Actuator. B Chem. 77, 401–408.
Guo, T., Yong, W., Jin, Y., Zhang, L., Liu, J., Wang, S., Chen, Q., Dong, Y., Su, H., Tan, T., 2017. Applications of DART-MS for food quality and safety assurance in food supply chain. Mass Spectrom. Rev. 36, 161–187.
Guyon, F., Gaillard, L., Salagoïty, M.-H., Médina, B., 2011. Intrinsic ratios of glucose, fructose, glycerol and ethanol 13C/12C isotopic ratio determined by HPLC-co-IRMS: toward determining constants for wine authentication. Anal. Bioanal. Chem. 401, 1551.
Heine, K., Phan, M., Atwal, G., 2016. Authenticity and prestige: what luxury brands could learn from the wine industry? Luxury Res. J. 1, 177–190.
Hellman, P., 2008. *Domaine Ponsot Proprietor Halts Sale of Fake Bottles* [Online]. Domaine Ponsot proprietor halts sale of fake bottles. Available: https://www.winespectator.com. (Accessed 20 May 2019).
Herbert, P., Barros, P., Alves, A., 2000. Detection of port wine imitations by discriminant analysis using free amino acids profiles. Am. J. Enol. Vitic. 51, 262–268.
Higgins, L.M., Llanos, E., 2015. A healthy indulgence? Wine consumers and the health benefits of wine. Wine Econ. Policy 4, 3–11.
Higgins, L.M., McGarry Wolf, M., Wolf, M.J., 2014. Technological change in the wine market? The role of QR codes and wine apps in consumer wine purchases. Wine Econ. Policy 3, 19–27.
Hjelmeland, A.K., King, E.S., Ebeler, S.E., Heymann, H., 2012. Characterizing the chemical and sensory profiles of United States Cabernet Sauvignon wines and blends. Am. J. Enol. Vitic. 64, 169–179.
Holmberg, 2010. Wine fraud. Int. J. Wine Res. 2, 105–113.
Hopfer, H., Heymann, H., 2014. Judging wine quality: do we need experts, consumers or trained panelists? Food Qual. Prefer. 32, 221–233.
Hopfer, H., Nelson, J., Ebeler, S.E., Heymann, H., 2015. Correlating wine quality indicators to chemical and sensory measurements. Molecules 20, 8453–8483.
Hu, Z., Cole, P.H., 2010. Bottle packaged wine product detection by UHF RFID systems. In: International Conference on Electromagnetics in Advanced Applications, pp. 301–304.
Huang, X.-Y., Jiang, Z.-T., Tan, J., Li, R., 2017. Geographical origin traceability of red wines based on chemometric classification via organic acid profile. J. Food Qual. 2017, 2038073.
Huang, Z., Ough, C.S., 1991. Amino acid profiles of commercial grape juices and wines. Am. J. Enol. Vitic. 42, 261–267.
Işçi, B., Yildirim, H.K., Altindişli, A., 2009. A review of the authentication of wine origin by molecular markers. J. Inst. Brew. 115, 259–264.
Jackson, R.S., 2014. Chemical constituents of grapes and wine. In: Jackson, R.S. (Ed.), Wine Science. Academic Press, San Diego, pp. 347–426.
Jaitz, L., Siegl, K., Eder, R., Rak, G., Abranko, L., Koellensperger, G., Hann, S., 2010. LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. Food Chem. 122, 366–372.
Jeffery, D.W., Wilkinson, K.L., 2014. Wine. In: Bamforth, C.W., Ward, R.E. (Eds.), The Oxford Handbook of Food Fermentations. Oxford University Press, New York, NY, USA, pp. 54–147.
Kallithraka, S., Arvanitoyannis, I.S., Kefalas, P., El-Zajouli, A., Soufleros, E., Psarra, E., 2001. Instrumental and sensory analysis of Greek wines; implementation of principal component analysis (PCA) for classification according to geographical origin. Food Chem. 73, 501–514.
Kallithraka, S., Mamalos, A., Makris, D.P., 2007. Differentiation of young red wines based on chemometrics of minor polyphenolic constituents. J. Agric. Food Chem. 55, 3233–3239.
Kamiloglu, S., 2019. Authenticity and traceability in beverages. Food Chem. 277, 12–24.
Kapferer, J.-N., 2008. The New Strategic Brand Management: Creating and Sustaining Brand Equity Long Term. Kogan Page Publishers, London, UK.
Kehagia, O., Chrysochou, P., Chryssochoidis, G., Krystallis, A., Linardakis, M., 2007. European consumers' perceptions, definitions and expectations of traceability and the importance of labels, and the differences in these perceptions by product type. Sociol. Rural. 47, 400–416.
Kok, E., Van der Spiegel, M., Prins, T., Manti, V., Groot, M.J., Bremer, M., Raamsdonk, L.v., Van der Fels-Klerx, H.J., Ruth, S., 2012. Traceability. In: Picó, Y. (Ed.), Chemical Analysis of Food: Techniques and Applications. Academic Press, USA, pp. 465–498.
Kokkinofta, R., Fotakis, C., Zervou, M., Zoumpoulakis, P., Savvidou, C., Poulli, K., Louka, C., Economidou, N., Tzioni, E., Damianou, K., Loupasaki, S., Kefalas, P., 2017. Isotopic and elemental authenticity markers: a case study on cypriot wines. Food Analyt. Methods 10, 3902–3913.
Košir, I.J., Lapornik, B., Andrenšek, S., Wondra, A.G., Vrhovšek, U., Kidrič, J., 2004. Identification of anthocyanins in wines by liquid chromatography, liquid chromatography-mass spectrometry and nuclear magnetic resonance. Anal. Chim. Acta 513, 277–282.
Koutroupi, E., Natos, D., Karelakis, C., 2015. Assessing exports market dynamics: the case of Greek wine exports. Proc. Econ. Finance 19, 184–192.
Kumšta, M., Pavloušek, P., Kárník, P., 2014. Use of anthocyanin profiles when differentiating individual varietal wines and terroirs. Food Technol. Biotechnol. 52, 383–390.
Lachenmeier, D., 2016. Advances in the detection of the adulteration of alcoholic beverages including unrecorded alcohol. In: Downey, G. (Ed.), Advances in Food Authenticity Testing. Woodhead Publishing, UK, pp. 565–584.
Larson, A., 2012. Generation Yine: The Millennial Generation and Wine Label Trends. Bachelor of Science Thesis, California Polytechnic State University.
Lecat, B., Brouard, J., Chapuis, C., 2017. Fraud and counterfeit wines in France: an overview and perspectives. Br. Food J. 119, 84–104.
Li, L., 2013. Technology designed to combat fakes in the global supply chain. Bus. Horiz. 56, 167–177.
Liao, S., Ma, Y.-Y., 2009. Conceptualizing consumer need for product authenticity. Int. J. Bus. Inf. 4, 89–114.
Liu, L., Cozzolino, D., Cynkar, W.U., Gishen, M., Colby, C.B., 2006. Geographic classification of Spanish and Australian Tempranillo red wines by visible and near-infrared spectroscopy combined with multivariate analysis. J. Agric. Food Chem. 54, 6754–6759.
Liu, Z., Wu, H.-L., Xie, L.-X., Hu, Y., Fang, H., Sun, X.-D., Wang, T., Xiao, R., Yu, R.-Q., 2017. Direct and interference-free determination of thirteen phenolic compounds in red wines using a chemometrics-assisted HPLC-DAD strategy for authentication of vintage year. Analyt. Methods 9, 3361–3374.
Lozano, J., Santos, J.P., Horrillo, M.C., 2016. Wine applications with electronic noses. In: Rodríguez Méndez, M.L. (Ed.), Electronic Noses and Tongues in Food Science. Academic Press, San Diego, pp. 137–148.
Lubinga, M., Ngqangweni, S., Nyhodo, B., Xolisiwe Potelwa, Y., van Walt, S., Phaleng, L., Ntshangase, T., 2017. Geographical Indication (GI) in the wine industry: does it matter? Natl. Agric. Mark. Counc. 1, 1–22.
Lunardo, R., Guerinet, R., 2007. The influence of label on wine consumption: its effects on young consumers' perception of authenticity and purchasing behavior. In: International Marketing and Trade of Quality Food Products, pp. 279–291.
Luykx, D.M.A.M., van Ruth, S.M., 2008. An overview of analytical methods for determining the geographical origin of food products. Food Chem. 107, 897–911.
Magdas, D.A., Cinta Pinzaru, S., Guyon, F., Feher, I., Cozar, B.I., 2018a. Application of SERS technique in white wines discrimination. Food Contr. 92, 30–36.
Magdas, D.A., Cuna, S., Cristea, G., Ionete, R.E., Costinel, D., 2012. Stable isotopes determination in some Romanian wines. Isot. Environ. Health Stud. 48, 345–353.
Magdas, D.A., Guyon, F., Feher, I., Pinzaru, S.C., 2018b. Wine discrimination based on chemometric analysis of untargeted markers using FT-Raman spectroscopy. Food Contr. 85, 385–391.
Makris, D.P., Stamatina, K., Mamalos, A., 2006. Differentiation of young red wines based on cultivar and geographical origin with application of chemometrics of principal polyphenolic constituents. Talanta 70, 1143–1152.
Mandrile, L., Zeppa, G., Giovannozzi, A.M., Rossi, A.M., 2016. Controlling protected designation of origin of wine by Raman spectroscopy. Food Chem. 211, 260–267.
Mannino, S., Benedetti, S., Buratti, S., Cosio, M.S., Scampicchio, M., 2007. Chapter 31 Electrochemical sensors for food authentication. Compr. Anal. Chem. 49, 755–770.
Margalit, Y., 2004. Must and wine composition. In: Concepts in Wine Chemistry. Wine Appreciation Guild, San Francisco, USA, pp. 1–59.
Márquez, R., Castro-Mejías, R., Natera-Marín, R., García-Barroso, C., 2008. Characterisation of the volatile fraction of Andalusian sweet wines. Eur. J. Res. Technol. 226, 1479.
Martin, A.E., Watling, R.J., Lee, G.S., 2012. The multi-element determination and regional discrimination of Australian wines. Food Chem. 133, 1081–1089.

McCulloch, J., McCarthy, P., Guru, S.M., Peng, W., Hugo, D., Terhorst, A., 2008. Wireless sensor network deployment for water use efficiency in irrigation. In: Proceedings of the Workshop on Real-World Wireless Sensor Networks. ACM, USA, pp. 46–50.

McGovern, P., Jalabadze, M., Batiuk, S., Callahan, M.P., Smith, K.E., Hall, G.R., Kvavadze, E., Maghradze, D., Rusishvili, N., Bouby, L., Failla, O., Cola, G., Mariani, L., Boaretto, E., Bacilieri, R., This, P., Wales, N., Lordkipanidze, D., 2017. Early neolithic wine of Georgia in the South Caucasus. PNAS 114, E10309–E10318.

McKellar, R.C., Vasantha Rupasinghe, H., Lu, X., Knight, K.P., 2005. The electronic nose as a tool for the classification of fruit and grape wines from different Ontario wineries. J. Sci. Food Agric. 85, 2391–2396.

Médina, B., Salagoïty, M.H., Guyon, F., Gaye, J., Hubert, P., Guillaume, F., 2013. Using new analytical approaches to verify the origin of wine. In: Brereton, P. (Ed.), New Analytical Approaches for Verifying the Origin of Food. Woodhead Publishing, Cambridge, UK, pp. 149–188.

Mendes-Ferreira, A., Barbosa, C., Lage, P., Mendes-Faia, A., 2011. The impact of nitrogen on yeast fermentation and wine quality. Ciência Téc. Vitiv. 26, 17–32.

Mendes, I., Sanchez, I., Franco-Duarte, R., Camarasa, C., Schuller, D., Dequin, S., Sousa, M.J., 2017. Integrating transcriptomics and metabolomics for the analysis of the aroma profiles of *Saccharomyces cerevisiae* strains from diverse origins. BMC Genom. 18, 455.

Model, P., 2005. Tamper-proof Closure/seal for Containers, Particularly Wine Bottles. US2005/0258129 A1.

Moe, T., 1998. Perspectives on traceability in food manufacture. Trends Food Sci. Technol. 9, 211–214.

Moulard, J., Babin, B.J., Griffin, M., 2015. How aspects of a wine's place affect consumers' authenticity perceptions and purchase intentions: the role of country of origin and technical terroir. Int. J. Wine Bus. Res. 27, 61–78.

Moulard, J.G., Raggio, R.D., Folse, J.A.G., 2016. Brand authenticity: testing the antecedents and outcomes of brand management's passion for its products. Psychol. Market. 33, 421–436.

Niimi, J., Boss, P.K., Bastian, S.E.P., 2018. Sensory profiling and quality assessment of research Cabernet Sauvignon and Chardonnay wines; quality discrimination depends on greater differences in multiple modalities. Food Res. Int. 106, 304–316.

Noraziah, A., Johari, M.S., Herawan, T., Sidek, R.M., Lee, H.C., Abdalla, A.N., 2011. Managing registration of new student intake using ORNSIS-Integrated barcode technology. Proc. Soc. Behav. Sci. 28, 144–150.

Nunes-Miranda, J.D., Santos, H.M., Reboiro-Jato, M., Fdez-Riverola, F., Igrejas, G., Lodeiro, C., Capelo, J.L., 2012. Direct matrix assisted laser desorption ionization mass spectrometry-based analysis of wine as a powerful tool for classification purposes. Talanta 91, 72–76.

Oliveri, P., Downey, G., 2013. Discriminant and class-modelling chemometric techniques for food PDO verification. In: De La Guardia, M., Gonzálvez, A. (Eds.), Comprehensive Analytical Chemistry. Elsevier, UK, pp. 317–338.

Oliveri, P., Simonetti, R., 2016. Chemometrics for food authenticity applications. In: Downey, G. (Ed.), Advances in Food Authenticity Testing. Woodhead Publishing, UK, pp. 701–728.

Ortea, I., O'Connor, G., Maquet, A., 2016. Review on proteomics for food authentication. J. Proteom. 147, 212–225.

Palade, L.M., Popa, M.E., 2014. Wine traceability and authenticity – a literature review. Sci. Bull. F Biotechnol. XVIII, 226–233.

Palmasi, M., Nguyen, A., Lee, K., O'Boyle, L., 2001. Tamper Evident Holographic Devices and Methods of Manufacture. US 6,214,443 B1.

Pascu, E., 2013. The authenticity and traceability of food – consumers protection form. Ann. Univ. Oradea Econ. Sci. 22, 658–662.

Pinu, F.R., 2018. Grape and wine metabolomics to develop new insights using untargeted and targeted approaches. Fermentation 4, 92.

Przyswa, E., 2014a. Protecting your wine. In: Wines and Vines.

Przyswa, E., 2014b. Choosing a fraud-prevention system. In: Wines and Vines.

Pustjens, A.M., Weesepoel, Y., van Ruth, S.M., 2016. Food fraud and authenticity: emerging issues and future trends. In: Leadley, C.E. (Ed.), Innovation and Future Trends in Food Manufacturing and Supply Chain Technologies. Elsevier, UK, pp. 3–20.

Qesja, B., Crouch, R., Quester, P., 2016. Reduced alcohol wine: product innovation and authenticity: the case of wine. Wine Vitic. J. 31, 67.

Rapeanu, G., Constanta, V., Bichescu, C., 2009. Possibilities to assess the wines authenticity. Innovat. Rom. Food Biotechnol. 5, 1–9.

Ray, J., Johan, B., 2007. Regional brand image and perceived wine quality: the consumer perspective. Int. J. Wine Bus. Res. 19, 276–297.

Rešetar, D., Marchetti-Deschmann, M., Allmaier, G., Katalinić, J.P., Kraljević Pavelić, S., 2016. Matrix assisted laser desorption ionization mass spectrometry linear time-of-flight method for white wine fingerprinting and classification. Food Contr. 64, 157–164.

Riovanto, R., Cynkar, W.U., Berzaghi, P., Cozzolino, D., 2011. Discrimination between Shiraz wines from different Australian regions: the role of spectroscopy and chemometrics. J. Agric. Food Chem. 59, 10356–10360.

Robinson, A.L., Adams, D.O., Boss, P.K., Heymann, H., Solomon, P.S., Trengove, R.D., 2012. Influence of geographic origin on the sensory characteristics and wine composition of *Vitis vinifera* cv. Cabernet Sauvignon wines from Australia. Am. J. Enol. Vitic. 63, 467–476.

Romano, A., Fischer, L., Herbig, J., Campbell-Sills, H., Coulon, J., Lucas, P., Cappellin, L., Biasioli, F., 2014. Wine analysis by FastGC proton-transfer reaction-time-of-flight-mass spectrometry. Int. J. Mass Spectrom. 369, 81–86.

Rose, B., 2019. What the hell is blockchain? Aust. N. Z. Grapegrow. Winemak. 36, 38–39.

Rössel, J., Schenk, P., Eppler, D., 2016. The emergence of authentic products: the transformation of wine journalism in Germany, 1947–2008. J. Consum. Cult. 18, 453–473.

Roullier-Gall, C., Witting, M., Gougeon, R.D., Schmitt-Kopplin, P., 2014. High precision mass measurements for wine metabolomics. Front. Chem. 2, 102.

Roullier-Gall, C., Witting, M., Tziotis, D., Ruf, A., Gougeon, R.D., Schmitt-Kopplin, P., 2015. Integrating analytical resolutions in non-targeted wine metabolomics. Tetrahedron 71, 2983–2990.

Rubert, J., Lacina, O., Fauhl-Hassek, C., Hajslova, J., 2014. Metabolic fingerprinting based on high-resolution tandem mass spectrometry: a reliable tool for wine authentication? Anal. Bioanal. Chem. 406, 6791–6803.

Rubert, J., Zachariasova, M., Hajslova, J., 2015. Advances in high-resolution mass spectrometry based on metabolomics studies for food – a review. Food Addit. Contam. 32, 1685–1708.

Sáenz-Navajas, M.-P., Campo, E., Sutan, A., Ballester, J., Valentin, D., 2013. Perception of wine quality according to extrinsic cues: the case of Burgundy wine consumers. Food Qual. Prefer. 27, 44–53.

Schamel, G., Anderson, K., 2003. Wine quality and varietal, regional and winery reputations: hedonic prices for Australia and New Zealand. Econ. Rec. 79, 357–369.

Schlesier, K., Fauhl-Hassek, C., Forina, M., Cotea, V., Kocsi, E., Schoula, R., van Jaarsveld, F., Wittkowski, R., 2009. Characterisation and determination of the geographical origin of wines. Part I: overview. Eur. Food Res. Technol. 230, 1–13.

Schueuermann, C., Bremer, P., Silcock, P., 2017. PTR-MS volatile profiling of pinot noir wines for the investigation of differences based on vineyard site. J. Mass Spectrom. 52, 625–631.

Sebastian, M., 2009. Renal toxicity. In: Gupta, R.C. (Ed.), Handbook of Toxicology of Chemical Warfare Agents. Academic Press, San Diego, USA, pp. 561–574.

Sébédio, J.L., Malpuech-Brugère, C., 2016. Implementation of foodomics in the food industry. In: Galanakis, C.M. (Ed.), Innovation Strategies in the Food Industry. Academic Press, USA, pp. 251–269.

Sémon, E., Arvisenet, G., Guichard, E., Le Quéré, J.-L., 2018. Modified proton transfer reaction mass spectrometry (PTR-MS) operating conditions for in vitro and in vivo analysis of wine aroma. J. Mass Spectrom. 53, 65–77.

Sidali, K.L., Hemmerling, S., 2014. Developing an authenticity model of traditional food specialties: does the self-concept of consumers matter? Br. Food J. 116, 1692–1709.

Sillani, S., Miccoli, A., Nassivera, F., 2017. Different preferences for wine communication. Wine Econ. Policy 6, 28–39.

Sjostrom, T., Corsi, A.M., Lockshin, L., 2014. Consumer perceptions of premium and luxury wine brands. Wine Vitic. J. 29, 68–70.

Śliwińska, M., Wiśniewska, P., Dymerski, T., Wardencki, W., Namieśnik, J., 2016. Advances in electronic noses and tongues for food authenticity testing. In: Downey, G. (Ed.), Advances in Food Authenticity Testing. Woodhead Publishing, UK, pp. 201–225.

Spielmann, N., Babin, B., 2011. Testing congruency effects between origin and producer for wines. Int. J. Wine Bus. Res. 23, 338–354.
Spielmann, N., Charters, S., 2013. The dimensions of authenticity in terroir products. Int. J. Wine Bus. Res. 25, 310–324.
Spitaler, R., Araghipour, N., Mikoviny, T., Wisthaler, A., Via, J.D., Märk, T.D., 2007. PTR-MS in enology: advances in analytics and data analysis. Int. J. Mass Spectrom. 266, 1–7.
Styles, O., 2009. Vin de Pays d'Oc in massive US fraud scandal [Online]. Vin de Pays d'Oc in massive US fraud scandal, UK. Available: https://www.decanter.com/. (Accessed 22 May 2019).
Suro, R., 1986. Italy Acting to End the Sale of Methanol-Tainted Wine [Online]. USA: Italy acting to end the sale of methanol-tainted wine. Available: https://www.nytimes.com. (Accessed 22 May 2019).
Taylor, P., Cullimore, I., Wrench, D., 2006. System and Method of RFID Data Tracking in Winemaking Process. United States patent application US 2006/0113374. A1 01/06/2006.
Taylor, W.S., 1989. Transparent Wine Bottle Closure and Cork, 4,812,317.
Teixeira dos Santos, C.A., Pascoa, R.N.M.J., Lopes, J.A., 2017a. A review on the application of vibrational spectroscopy in the wine industry: from soil to bottle. Trends Anal. Chem. 88, 100–118.
Teixeira dos Santos, C.A., Pascoa, R.N.M.J., Sarraguca, M.C., Porto, P., Cerdeira, A.L., Gonzalez-Saiz, J.M., Pizarro, C., Lopes, J.A., 2017b. Merging vibrational spectroscopic data for wine classification according to the geographic origin. Food Res. Int. 102, 504–510.
Thomas, A., Pickering, G., 2003. The importance of wine label information. Int. J. Wine Mark. 15, 58–74.
Trygg, J., Gullberg, J., Johansson, A.I., Jonsson, P., Moritz, T., 2006. Chemometrics in metabolomics — an introduction. In: Saito, K., Dixon, R.A., Willmitzer, L. (Eds.), Plant Metabolomics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 117–128.
Tustin, M., Lockshin, L., 2001. Region of origin: does it really count. Aust. N. Z. Wine Ind. J. 16, 139–143.
Tzoulis, I., Andreopoulou, Z., 2013. Emerging traceability technologies as a tool for quality wood trade. Proc. Technol. 8, 606–611.
UNESCO, Association Terroirs & Cultures, 2005. Rencontres internationales planete terroirs. In: Planete Terroirs.
Unwin, T., 2012. Terroir: at the heart of geography. In: Dougherty, P.H. (Ed.), The Geography of Wine: Regions, Terroir and Techniques. Springer, Dordrecht, Netherlands, pp. 37–48.
Vaclavik, L., Lacina, O., Hajslova, J., Zweigenbaum, J., 2011. 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. Anal. Chim. Acta 685, 45–51.
Van Emon, J.M., 2016. The omics revolution in agricultural research. J. Agric. Food Chem. 64, 36–44.
Vaudour, E., 2002. The quality of grapes and wine in relation to geography: notions of terroir at various scales. J. Wine Res. 13, 117–141.
Verdú, J.A.J., Lloréns Montes, F.J., Fuentes Fuentes, M.M., 2004. Measuring perceptions of quality in food products: the case of red wine. Food Qual. Prefer. 15, 453–469.
Vergara, C., von Baer, D., Mardones, C., Gutiérrez, L., 2011. Overview of chemical markers for varietal authentication of red wines. In: Ebeler, S.E., Takeoka, G.R., Winterhalter, P. (Eds.), Progress in Authentication of Food and Wine. American Chemical Society, Washington, DC, USA, pp. 101–111.
Versari, A., Laurie, V.F., Ricci, A., Laghi, L., Parpinello, G.P., 2014. Progress in authentication, typification and traceability of grapes and wines by chemometric approaches. Food Res. Int. 60, 2–18.
Vilanova, M., Oliveira, J., 2012. Application of gas chromatography on the evaluation of grape and wine aroma in Atlantic viticulture (NW Iberian Peninsula). In: Salih, B., Çelikbıçak, Ö. (Eds.), Gas Chromatography in Plant Science, Wine Technology, Toxicology and Some Specific Applications. IntechOpen, Rijeka, Croatia, pp. 109–146. https://doi.org/10.5772/31854.
Villamor, R.R., Ross, C.F., 2013. Wine matrix compounds affect perception of wine aromas. Ann. Rev. Food Sci. Technol. 4, 1–20.
Villano, C., Lisanti, M.T., Gambuti, A., Vecchio, R., Moio, L., Frusciante, L., Aversano, R., Carputo, D., 2017. Wine varietal authentication based on phenolics, volatiles and DNA markers: state of the art, perspectives and drawbacks. Food Contr. 80, 1–10.
Vukatana, K., Sevrani, K., Hoxha, E., 2016. Wine traceability: a data model and prototype in Albanian context. Foods 5, 11.
Wang, L., Kowk, S.K., Ip, W.H., 2012. A radio frequency identification-based quality evaluation system design for the wine industry. Int. J. Comput. Integrated Manuf. 25, 11–19.
Want, R., 2006. An introduction to RFID technology. IEEE Pervasive Comput. 5, 25–33.
Waterhouse, A.L., 2002. Wine phenolics. Ann. N. Y. Acad. Sci. 957, 21–36.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016a. Authentication. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. 383–391.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016b. Introduction. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. xvii–xxiv.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016c. Water and ethanol. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. 1–10.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016d. Introduction to phenolics. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. 99–104.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016e. Minerals. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. 34–39.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016f. Introduction to phenolics. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. 99–104.
Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., 2016g. Grape must composition overview. In: Understanding Wine Chemistry. John Wiley & Sons, UK, pp. 172–178.
Weiss, J., Takhistov, P., McClements, D.J., 2006. Functional materials in food nanotechnology. J. Food Sci. 71, 107–116.
Welke, J.E., Manfroi, V., Zanus, M., Lazzarotto, M., Alcaraz Zini, C., 2013. Differentiation of wines according to grape variety using multivariate analysis of comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection data. Food Chem. 141, 3897–3905.
Westad, F., Marini, F., 2015. Validation of chemometric models – a tutorial. Anal. Chim. Acta 893, 14–24.
Wine Australia, 2018. Wine Australia act 2013. In: Agriculture and Water Resources.
Wine Traceability Working Group, 2018. *Wine Supply Chain Traceability GS1 Application Guideline* [Online]. Wine supply chain traceability GS1 application guideline. Available: https://www.gs1us.org/DesktopModules/Bring2mind/DMX/Download.aspx?Command=Core_Download&EntryId=660. (Accessed 7 May 2020).
Wongprawmas, R., Spadoni, R., 2018. Is innovation needed in the Old World wine market? The perception of Italian stakeholders. Br. Food J. 120, 1315–1329.
Wynn, M.T., Ouyang, C., ter Hofstede, A.H.M., Fidge, C.J., 2011. Data and process requirements for product recall coordination. Comput. Ind. 62, 776–786.
Xi, J., Ye, T.T., 2012. Conformal UHF RFID tag antenna mountable on winebottle neck. In: Proceedings of the 2012 IEEE International Symposium on Antennas and Propagation. Institute of Electrical and Electronic Engineers, USA, pp. 1–2.
Yick, J., Mukherjee, B., Ghosal, D., 2008. Wireless sensor network survey. Comput. Network. 52, 2292–2330.
Yiu, N.C.K., 2014. An NFC-Enabled Anti-counterfeiting System for Wine Industry. PhD Thesis. The University of Hong Kong.
Youssef, S.M., Salem, R.M., 2007. Automated barcode recognition for smart identification and inspection automation. Expert Syst. Appl. 33, 968–977.
Zhang, J., Li, L., Gao, N., Wang, D., Gao, Q., Jiang, S., 2010. Feature extraction and selection from volatile compounds for analytical classification of Chinese red wines from different varieties. Anal. Chim. Acta 662, 137–142.
Zhao, X., Procopio, S., Becker, T., 2015. Flavor impacts of glycerol in the processing of yeast fermented beverages: a review. J. Food Sci. Technol. 52, 7588–7598.

Chapter 2

Review Article & Research Aims

Publication

A Review of Wine Authentication Using Spectroscopic Approaches in Combination with Chemometrics.

Ranaweera, R.K.R.; Capone, D.L.; Bastian, S.E.P.; Cozzolino, D.; Jeffery, D.W. A review of wine authentication using spectroscopic approaches in combination with chemometrics. Molecules 2021, 26, 4334.

Statement of Authorship

Title of Paper	A Review of Wine Authentication Using Spectroscopic Approaches in Combination with Chemometrics
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Ranaweera, R.K.R.; Capone, D.L.; Bastian, S.E.P.; Cozzolino, D.; Jeffery, D.W. A Review of Wine Authentication Using Spectroscopic Approaches in Combination with Chemometrics. *Molecules* **2021**, *26*, 4334. https://doi.org/10.3390/molecules26144334

Principal Author

Name of Principal Author (Candidate)	Ranaweera K. R. Ranaweera		
Contribution to the Paper	Contributed to conceptualisation, investigation of literature, visualisation of data, and writing the first manuscript draft, along with the reviewing and editing of the final manuscript.		
Overall percentage (%)	75%		
Certification:	This paper is a review on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/07/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dimitra L. Capone		
Contribution to the Paper	Conceptualisation, supervision, critical review and editing.		
Signature		Date	20/07/2021

Name of Co-Author	Susan E.P. Bastian		
Contribution to the Paper	Conceptualisation, supervision, critical review and editing.		
Signature		Date	20·7·2021

Name of Co-Author	Daniel Cozzolino		
Contribution to the Paper	Conceptualisation, supervision, critical review and editing.		
Signature		Date	3/08/2021

Name of Co-Author	David W. Jeffery		
Contribution to the Paper	Conceptualisation, investigation of literature, funding acquisition, reviewed, edited and revised the manuscript for publication. Acted as the corresponding author at all stages.		
Signature		Date	19/07/20201

molecules

MDPI

Review

A Review of Wine Authentication Using Spectroscopic Approaches in Combination with Chemometrics

Ranaweera K. R. Ranaweera [1], Dimitra L. Capone [1,2], Susan E. P. Bastian [1,2], Daniel Cozzolino [3] and David W. Jeffery [1,2,*]

[1] Department of Wine Science and Waite Research Institute, The University of Adelaide, PMB 1, Glen Osmond, SA 5064, Australia; ruchira.ranaweera01@adelaide.edu.au (R.K.R.R.); dimitra.capone@adelaide.edu.au (D.L.C.); sue.bastian@adelaide.edu.au (S.E.P.B.)
[2] Australian Research Council Training Centre for Innovative Wine Production, The University of Adelaide, PMB 1, Glen Osmond, SA 5064, Australia
[3] Queensland Alliance for Agriculture and Food Innovation (QAAFI), The University of Queensland, Hartley Teakle Building, Brisbane, QLD 4072, Australia; d.cozzolino@uq.edu.au
* Correspondence: david.jeffery@adelaide.edu.au; Tel.: +61-8-8313-6649

Abstract: In a global context where trading of wines involves considerable economic value, the requirement to guarantee wine authenticity can never be underestimated. With the ever-increasing advancements in analytical platforms, research into spectroscopic methods is thriving as they offer a powerful tool for rapid wine authentication. In particular, spectroscopic techniques have been identified as a user-friendly and economical alternative to traditional analyses involving more complex instrumentation that may not readily be deployable in an industry setting. Chemometrics plays an indispensable role in the interpretation and modelling of spectral data and is frequently used in conjunction with spectroscopy for sample classification. Considering the variety of available techniques under the banner of spectroscopy, this review aims to provide an update on the most popular spectroscopic approaches and chemometric data analysis procedures that are applicable to wine authentication.

Keywords: authenticity; multivariate analysis; wine fingerprinting; spectral data; machine learning

Citation: Ranaweera, R.K.R.; Capone, D.L.; Bastian, S.E.P.; Cozzolino, D.; Jeffery, D.W. A Review of Wine Authentication Using Spectroscopic Approaches in Combination with Chemometrics. *Molecules* **2021**, *26*, 4334. https://doi.org/10.3390/molecules26144334

Academic Editor: Angelo Antonio D'Archivio

Received: 15 June 2021
Accepted: 14 July 2021
Published: 17 July 2021

1. Introduction

Wine is a historic alcoholic beverage that has evolved to be of high commercial importance. It can be identified as a luxurious commodity and is produced and consumed in many countries around the world. Wine consists of innumerable compounds spanning various concentration ranges, many of which are essential to its evolution and quality, as well as for human health benefits in the case of red wine [1]. In general, the composition of red wine can be broadly represented on a w/w basis as 86% water, 11% ethanol, and 3% for the remainder, which includes glycerol, sugars, polyols, phenols, minerals, organic acids, and volatile compounds [2]. The composition of wine mainly depends on certain factors that define the wine's identity, including grape variety, geographical origin, the biophysical environment of the vineyard, vintage conditions, and winemaking inputs [3]. Different types of fraud related to those factors have been encountered in wine over the years, including counterfeiting of labels and brands, adulteration through the use of unauthorised additives or practices, and substitution based on grape variety or region of origin [4]. Therefore, to confirm the genuineness of wine and protect its value, analytical techniques need to be applied to explore the chemical constituents of wine that aid in the development of models for authenticity.

Classical techniques such as gas chromatography-mass spectrometry and high-performance liquid chromatography are advancing continuously, facilitating wine analysis with high sensitivity [5]. Considering the applicability in an industrial setting, however, aspects such as rapidity, user-friendliness, and cost-effectiveness have become of paramount

importance in recent times [6]. Spectroscopic techniques provide a great solution due to their relative simplicity, speed of analysis, simple sample preparation, and environmental friendliness [7], and have been well-utilised for different wine and grape research studies, such as for targeted and non-targeted chemical analyses [8], prediction of sensory attributes [9], and wine authentication [10,11]. A snapshot of selected research outcomes identified from the Web of Science Core Collection over the past three decades using 'wine authentication' and 'spectroscopy' as the search keywords is visualised in Figure 1 to provide some understanding of the trends in the literature. Aside from those specific keywords, the terms classification, chemometrics, and geographical origin also feature prominently and are variously linked upon closer inspection to a range of terms associated with spectroscopic (e.g., near-infrared, mid-infrared, NMR, UV–visible, Raman, fluorescence) and chemometric (e.g., partial least squares discriminant analysis, feature selection, support vector machines, artificial neural networks, principal component analysis, discriminant analysis, data fusion, pattern recognition) techniques.

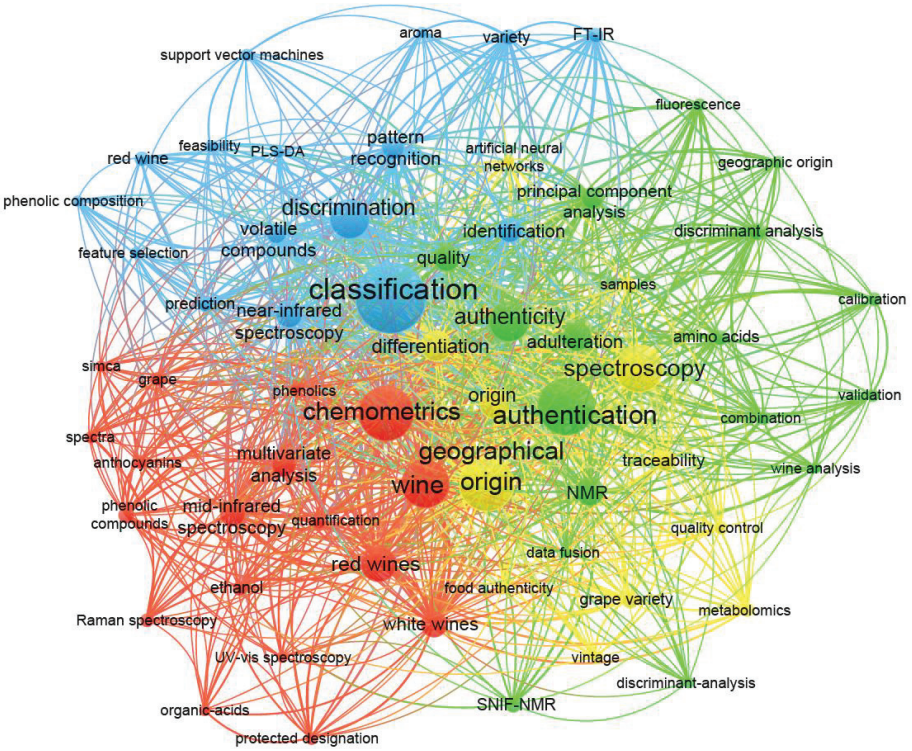


Figure 1. Bibliometric map of wine science-related research visualised from 222 publications (from 1990 to 2021) recovered from Web of Science Core Collection using 'wine authentication' and 'spectroscopy' as keywords. Literature analysis and figure construction were facilitated with VOSviewer [12]. Different colours are used to define the clusters that terms belong to. The bibliometric relationship between terms is indicated using curved lines and the relative size of the words reflects the number of publications in which the terms occurred.

Among the different spectroscopic methods that are available, techniques such as nuclear magnetic resonance (NMR), near-infrared (NIR), mid-infrared (MIR), Raman, and fluorescence have been prominent in past research studies. Moreover, it is clear from Figure 1 that chemometric techniques (i.e., multivariate data analyses) have been an integral part of these spectroscopic techniques to draw meaningful conclusions regarding sample classification and differentiation. Taking these aspects together, this review emphasises the application of spectroscopic techniques and chemometrics to authenticity

Molecules **2021**, *26*, 4334

in the field of wine research using examples from the past 15 years. The strengths and weaknesses of different spectroscopic methods for wine authentication are presented and various chemometric methods applied to address specific requirements in classification are discussed. Finally, future trends and directions for wine authentication with spectroscopic approaches have been identified.

2. Spectroscopic Techniques Applied in Wine Authentication

Wine authentication verifies that the label description is in compliance with the content of the package through an analytical process [13], which can be carried out through targeted or non-targeted methods. In targeted analyses, variations of a specific marker compound or certain metabolites are considered for differentiation of samples, whereas in non-targeted analyses, a chemical 'fingerprint' of the sample is obtained and similarities/differences in fingerprint are used for classification with the aid of chemometrics [14]. Spectroscopic techniques are frequently utilised for non-targeted wine fingerprinting.

In spectroscopic analysis, chemical and physical (structural) information within samples is exploited according to the interaction of atoms and molecules with electromagnetic radiation (Figure 2), which related to the wavelength or frequency spectrum of either absorbed or emitted energy [14]. For instance, ultraviolet-visible (UV–Vis) absorption and fluorescence spectroscopy is based on changes that occur in electronic states. In another way, infrared (IR) and Raman spectroscopic techniques are based on vibrational variations in the molecules. Moving further along the electromagnetic spectrum to longer wavelengths past the microwave region, NMR involves changes in rotational state, with nuclear spin being affected within the radiofrequency range. Data obtained from these methods typically needs to be analysed through multivariate techniques to obtain useful information hidden in the spectra. For authentication purposes, data can then be further classified using statistical approaches as outlined in Section 3. Firstly though, the main spectroscopic techniques applied for wine classification (as identified from Figure 1) are reviewed, which necessarily involves some mention of chemometrics.

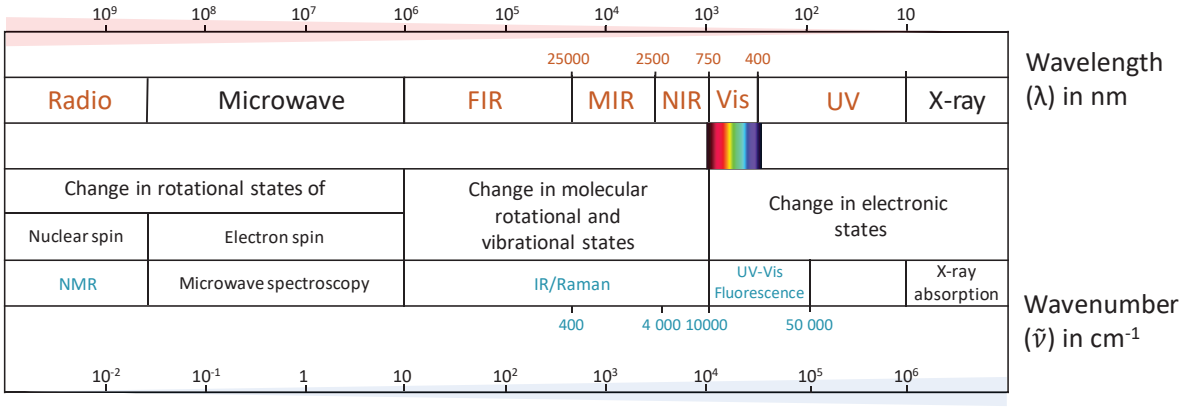


Figure 2. The electromagnetic spectrum and its relevance to different spectroscopic methods. FIR, far-infrared; MIR, mid-infrared; NIR, near-infrared; Vis, visible; UV, ultraviolet. Techniques in this review that are applied for wine authentication are indicated in light blue font. Conceptualised from [15].

2.1. UV–Vis Spectroscopy

UV–Vis spectroscopy is a fast, low-cost, and reliable analytical method that has been used in the analysis of wine for many decades [16]. Spectra recorded at UV and visible wavelengths (typically 190–800 nm, Figure 2) provide information about compounds in wine containing a chromophore, such as hydroxybenzoic (280 nm) and hydroxycinnamic (320 nm) acids, flavan-3-ols (280 nm), flavonols (370 nm), and anthocyanin glucosides (520 nm) [17]. As summarised in Table 1, UV–Vis spectroscopy has been applied in

wine discrimination according to the region of origin [18,19], grape variety and ageing process [20,21]. Although the specific chemical markers are not necessarily identified, as a non-targeted method combined with appropriate chemometric techniques such as linear discriminant analysis (LDA) and partial least squares discriminant analysis (PLS-DA), Azcarate et al. were able to correctly classify Argentinian Sauvignon blanc wine samples with 100% accuracy according to their geographical origin [19]. In their study, Philippidis et al. achieved 97.5% correct classification of grape variety and showed that the latent variables resulting from orthogonal projections to latent structures-discriminant analysis (OPLS-DA) could be related to the absorption of aromatic compounds such as phenolic acids and flavonols [21]. In comparison to other spectroscopic methods, however, UV–Vis spectroscopy provides a limited number of spectral features; therefore, it could be used as a screening approach with more sophisticated techniques being implemented for further analysis. In addition, the combination of other spectroscopic methods like IR and fluorescence with UV–Vis spectroscopy can improve the accuracy of classification models used for authentication by fusion of the datasets [22,23].

Table 1. Examples of UV–visible spectroscopy in combination with chemometrics for wine authentication.

Spectroscopic Technique	Spectral Range	Parameters for Authentication	Classification Method [1]	Remark	Reference
UV–Vis	200–800 nm	Geographical origin (Spanish denomination of origin)	SVM	Correct classification rates above 96%	[18]
UV–Vis	200–500 nm	Geographical origin of Argentinian regions	PCA, LDA, PLS-DA	Correct classification with LDA and PLS-DA methods of 100%	[19]
UV–Vis	300–800 nm	Discrimination by origin, grape variety and ageing process	PCA, SIMCA	Correct classification of 90% for geographical origin, and 75% for variety and ageing process	[20]
UV–Vis	240–700 nm	Discrimination according to grape variety, ageing process and barrel/container type	OPLS-DA	Correct classification of 97% for variety, 73% for ageing process and 100% for container type	[21]

[1] SVM, support vector machine; PCA, principal component analysis; LDA, linear discriminant analysis; PLS-DA, partial least squares-discriminant analysis; SIMCA, soft independent modelling of class analogy; OPLS-DA, orthogonal projections to latent structures discriminant analysis.

2.2. IR Spectroscopy

IR spectroscopy has been used in wine analysis for several decades [24] and has become the most frequently applied spectroscopic technique in comparison to other methods [25]. It is a user-friendly and rapid technique that provides information on many components in a wine matrix, and can be used for determination of parameters such as alcohol content, pH, volatile acidity, organic acids, reducing sugars, and polyphenols [26]. Two main IR-based techniques are applied according to the range in the spectral region: near-infrared (NIR) from 14,000 to 4000 cm^{-1} and mid-infrared (MIR) from approximately 4000 to 400 cm^{-1} (Figure 2). NIR spectra contain less intense bands than MIR and it is difficult to assign chemical groups specifically with NIR due to overlapping signals with water and ethanol around 1950 nm. In MIR, there is a 'fingerprint region' (1500–400 cm^{-1}) that includes unique absorption patterns of compounds such as phenolics that are mainly

applied for discrimination purposes, and signals associated with various functional groups can be assigned, such as C=O related to organic acids at 1700 cm^{-1}, and combinations of C–H vibrations and overtones related to ethanol and sugars at around 2300–2100 cm^{-1} [27]. Applicability of IR methods to wine analysis increased with the introduction of techniques such as Fourier transform (FT), which has improved data collection speed and reproducibility [28], and the application of attenuated total reflectance (ATR), which simplifies the sample handling process and is advantageous in routine analysis [29]. Classification of wine with IR has often been complemented by the use of UV and/or visible spectroscopy to enhance the accuracy of the classification [30]. Table 2 includes some examples of the application of IR spectroscopy (with or without UV–Vis) to wine authentication, along with the spectral region and classification method used.

Table 2. Examples of IR spectroscopy in combination with chemometrics for wine authentication.

Spectroscopic Technique	Spectral Range	Parameters for Authentication	Classification Method [1]	Remark	Reference
MIR	5012–926 cm^{-1}	Discrimination of red and white varieties from Australian regions	PCA, LDA	Correct classification of red varieties, 96% and white varieties, 94%	[26]
UV–Vis, NIR and MIR	400–2500 nm (UV–Vis and NIR) and 4000–400 cm^{-1} (MIR)	Geographical origin of Sauvignon blanc wines from Australia and New Zealand	PCA, SIMCA, PLS-DA	Correct classification using PLS-DA with: UV–Vis, 67%; NIR, 76%; MIR, 90%; and combined IR spectra, 93%	[23]
UV–Vis/NIR	190–2500 nm	Discrimination of white wines (Albariño cultivar) from Rías Baixas subzones in Spain	PCA, LDA, SIMCA, SVM	Correct classification using: LDA, 86%; SIMCA, 56%; and SVM, 84%	[30]
NIR and MIR	1750–1000 cm^{-1} and 4555–4353 cm^{-1}	Geographical origin of Cabernet Sauvignon wines from Australia, Chile, and China	PCA, SIMCA, DA	Correct classification using: SIMCA, 97%, 97%, and 92% for Australian, Chilean, and Chinese wines; and DA, 86%, 85%, and 77%, respectively.	[31]

[1] PCA, principal component analysis; LDA, linear discriminant analysis; SIMCA, soft independent modelling of class analogy; PLS-DA, partial least squares-discriminant analysis; SVM, support vector machine.

In another study, Bevin et al. discriminated Australian red wine (Cabernet Sauvignon, Shiraz and Merlot) and white wine (Chardonnay, Riesling, Sauvignon blanc and Viognier) according to grape variety with 96% and 94% accuracy, respectively, using LDA with MIR spectra [26]. Although subtle variation in wine composition contributed to these varietal discriminations, MIR signals are highly sensitive to temperature and pH, which needs to be considered in the application. For geographical authentication, Cozzolino et al. combined NIR and MIR techniques for Sauvignon blanc wines from Australia and New Zealand, achieving an overall 93% correct classification with PLS-DA, which was higher than for the individual IR techniques or for UV–Vis [23]. Similarly, the feasibility of differentiating subzones within a denomination of origin (DO) has been evaluated by Martelo-Vidal et al., who achieved their highest overall correct classification of 86% with LDA in comparison to soft independent modelling of class analogy (SIMCA, 56%) and support vector machine (SVM, 84%) for combined UV–Vis and NIR spectra [30]. Hu et al. applied MIR and NIR to classify Cabernet Sauvignon wines with SIMCA and correctly classified Australian, Chilean, and Chinese wines with 97%, 97%, and 92% accuracy, respectively [31]. Although these works yielded an accuracy of > 90% for classification, IR spectroscopy has limitations in quantitative analysis when measuring low abundance components (<0.5 g L^{-1}) [32].

2.3. Raman Spectroscopy

In comparison to other spectroscopic techniques, Raman spectroscopy has not been exploited much for wine analysis until recently [33]. This spectroscopic method involves detecting the inelastic scattered light emitted from molecular vibrations of a sample, approximately in the range 200–3600 nm (Figure 2). The Raman effect produces a weak signal, but the development of optimised detection capability provides the opportunity to obtain rich information regarding the chemical composition and dynamics of the sample [34]. There are two different regions in Raman spectroscopy, with Stokes Raman scattering having more dominant ethanol, sucrose and water peaks, and anti-Stokes Raman scattering from minor components such as aromatic compounds, including various phenolics, which can be more applicable to wine discrimination [35]. Indeed, for analysis of water dominant samples such as wine, Raman spectroscopy has an advantage over IR techniques because of the relatively weak signals from water molecules in the vibrational fingerprint range [36]. Two types of Raman technique are applied in food analysis: FT-Raman spectroscopy and surface-enhanced Raman spectroscopy (SERS). Both of these methods have been developed for the purpose of wine authentication, as shown in Table 3.

Table 3. Examples of Raman spectroscopy in combination with chemometrics for wine authentication.

Spectroscopic Technique	Spectral Region	Parameters for Authentication	Classification Method [1]	Remark	Reference
FT-Raman	1700–0 cm^{-1} (Stokes), −1000–0 cm^{-1} (anti-Stokes) (laser emitting at 1064 nm)	Discrimination of wines geographically, varietally, and by vintage	LDA	Correct classification of: variety, 84%; geographical origin, 100%; and vintage, 95%	[37]
SERS	3350–200 cm^{-1} (laser emitting at 532 nm)	Discrimination of wines geographically (Romanian and French and different Romanian regions), varietally, and by vintage	LDA	Correct classification of: variety, 90%; geographical origin, 83% among Romanian wines and 100% between countries; and vintage, 90%	[35]
SERS	1600–450 cm^{-1} (laser emitting at 785 nm)	Discrimination of wines according to variety and producer	PCA, SIMCA	Correct classification of: variety, 87%; and producer, 93%	[38]

[1] LDA, linear discriminant analysis; PCA, principal component analysis; SIMCA, soft independent modelling of class analogy.

The effectiveness of FT-Raman was shown in the work of Magdas et al., who discriminated white wine according to variety (Sauvignon, Riesling, Chardonnay, Pinot Gris), geographical origin (Romania and France), and vintage using LDA, achieving overall correct classification of 84%, 100%, and 95%, respectively [34,37]. In another study, Magdas and colleagues applied SERS to discriminate among white wines and compared it with FT-Raman, identifying a few common marker compounds between the techniques, such as ferulic and sinapic acids that resulted in differences among the wines. SERS was able to enhance the signals of more minor compounds such as caffeic acid, *p*-coumaric acid and resveratrol [35]. Applying the same SERS approach, Zanuttin et al. discriminated wines according to variety and producer with SIMCA, deriving an overall correct classification of 87%. Moreover, they identified major metabolites such as purines, carboxylic acids and glutathione that can be assigned to specific bands responsible for discrimination of wine [38]. The advantage of SERS over FT-Raman is the selectivity afforded by specific molecules being adsorbed to metal nanostructures (mainly noble metals), which enhances the intensity of Raman signals in SERS [39]. Complexity arises with the sample preparation step, however, as it is necessary to prepare a colloidal dispersion of Ag nanoparticles to add to the sample, which can be a disadvantage. Raman spectroscopy

requires spectral pre-processing such as multiplicative scatter correction (discussed in Section 3) to avoid the effect of fluorescence that can obscure Raman scattering, especially when analysing wine [40].

2.4. Fluorescence Spectroscopy

Fluorescence spectroscopy has been deemed as a useful tool in wine authentication for some time and its application has been enhanced recently with improvements in the chemometric analysis [41]. Because of the high sensitivity, selectivity, and rapidity of the technique, fluorescence spectroscopy has an advantage as an analytical platform [42]. It is based on the emission of longer wavelength light from a substance after absorption of energy in the UV or visible range (as with UV–Vis spectroscopy, Figure 2). Fluorescence typically occurs for aromatic molecules and can be well applied to wine analysis, with common fluorophores being a variety of phenolic compounds, vitamins, and aromatic amino acids [43]. According to the fluorophoric molecular and macromolecular constituents in the sample, a three-dimensional excitation-emission matrix (EEM) recorded over multiple excitation and emission wavelengths can be obtained and considered as the 'molecular fingerprint' of the sample [44]. Therefore, this approach in combination with chemometrics can be utilised for authentication of wine. When undertaking spectrofluorometric analysis, it is important to apply corrections for Rayleigh masking, Raman scattering, and inner filter effects (IFE), as well as to maintain proper pH and temperature to avoid the consequence of quenching, which can affect the fluorescence intensity. Several types of fluorescence methods can be applied to wine analysis according to the manner of obtaining the spectrum (i.e., total luminescence spectroscopy yielding an EEM or synchronous fluorescence spectroscopy) and by the geometry of sample illumination (i.e., right-angle for diluted samples or front-face for bulk liquids or solids) [45]. Fluorescence spectroscopy has been applied in several studies recently, in combination with chemometric techniques, for discrimination of wine according to geographical origin or variety (Table 4).

Table 4. Examples of fluorescence spectroscopy in combination with chemometrics for wine authentication.

Spectroscopic Technique	Spectral Region	Parameters for Authentication	Classification Method [1]	Remark	Reference
Synchronous fluorescence spectroscopy	λ_{ex} = 250–350 nm and λ_{em} = 250–500 nm	Discrimination of white wines according to variety in Tokaj (Slovakia)	PCA, LDA	Correct classification of variety, 100%	[46]
Total fluorescence spectroscopy	EEM λ_{ex} =240–800 nm and λ_{em} 242–824 nm	Discrimination of Cabernet Sauvignon wines from Australia and Bordeaux, France	SVMDA XGBDA	Correct classification of geographical origin using: XGBDA, 100%; and SVMDA, 85%	[11]
Total fluorescence spectroscopy	EEM λ_{ex} = 250–500 nm and λ_{em} 275–600 nm	Discrimination of white wine from Romania and France for geographical origin and variety	PARAFAC, SIMCA	Correct classification of: variety, 97%; and geographical origin, 98%	[47]
Total fluorescence spectroscopy	EEM λ_{ex} =240–700 nm and λ_{em} 242–824 nm	Discrimination of red wine varieties from different Australian regions for variety and geographical origin	XGBDA	Correct classification of: variety, 100%; and geographical origin, 99.7%	[48]

[1] PCA, principal component analysis; LDA, linear discriminant analysis; SVMDA, support vector machine discriminant analysis; XGBDA, extreme gradient boosting discriminant analysis; PARAFAC, parallel factor analysis; SIMCA, soft independent modelling of class analogy.

Sádecká and Jakubíková applied synchronous fluorescence spectroscopy to discriminate white wine according to variety (Furmint, Lipovina, and Muscat blanc) using LDA, achieving an overall rate of 100% correct classification in validation and 93% for prediction [46]. Using total luminescence spectroscopy for authentication, Suciu et al. classified

white wine according to geographical origin (Romania and France) and variety (Chardonnay, Pinot Gris, Riesling and Sauvignon), obtaining correct classification rates of 98% and 97.1%, respectively, by applying parallel factor analysis (PARAFAC) and SIMCA algorithms [47]. Based on an absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) approach that also uses right-angle geometry with total fluorescence spectroscopy, Ranaweera et al. classified Cabernet Sauvignon wines according to geographical origin with 100% accuracy using EEM data and a machine learning algorithm known as extreme gradient boosting discriminant analysis (XGBDA) [11]. This was contrasted with SVM as an alternative machine learning technique, which gave 85% correct classification according to region. In a subsequent study, those authors used A-TEEM in conjunction with XGBDA to classify over 200 commercially produced but unreleased Australian red wines by origin and variety with 99.7% and 100% accuracy, respectively [48]. This method involved multi-block data analysis of EEM and absorbance datasets, as well as PARAFAC to extract components according to the major fluorophores differentiating the wines.

2.5. NMR Spectroscopy

Among the most mature forms of spectroscopy for food and beverage classification, NMR has been applied to wine authentication for decades. Initially, site-specific natural isotopic fractionation NMR (SNIF-NMR) spectroscopy was proposed as a tool for detecting the biochemical origin of ethanol according to the natural distribution of deuterium [49], which can reveal the unauthorised use of chaptalisation (sugar addition) in winemaking, for example [50]. NMR spectroscopy can be applied for qualitative analysis to determine molecular structures and for compositional profiling of a sample [51], as well as for quantitative analysis of analytes such as amino acids, alcohols, sugars, carboxylic acids and their derivatives, and phenolic compounds [50]. NMR can be based on acquisition of ^{1}H, ^{2}H, or ^{13}C spectra; for wine authentication, ^{1}H NMR spectroscopy is most advantageous as data acquisition is fast and highly reproducible compared to other techniques [33]. Moreover, NMR with advancements such as automation of analysis has been introduced commercially and adapted to wine authentication (e.g., Bruker's WineScreener^{TM}) [50]. Using the possibilities of NMR spectroscopy, different aspects of wine authenticity have been addressed (Table 5).

Table 5. Examples of NMR spectroscopy in combination with chemometrics for wine authentication.

Spectroscopic Technique	Spectral Range	Parameters for Authentication	Classification Method [1]	Remark	Reference
^{1}H NMR	0.5–9.5 ppm	Discrimination of wines geographically (German wine regions), varietally, and by vintage	PCA, LDA, NCM	Correct classification of: variety, 95%; geographical origin, 89%; and vintage, 96–97%	[52]
^{1}H NMR	0.8–9.7 ppm	Varietal differentiation of red and white wines produced in different regions in China	PCA, LDA	Correct classification of: red wines, 83%; and white wines, 94%	[53]
^{1}H NMR	0.0–10.0 ppm	Varietal differentiation of red and white wines produced in Czech Republic	PCA, RF	Correct classification of: most varieties, ~70%; and type of wine, 92%	[54]

[1] PCA, principal component analysis; LDA, linear discriminant analysis; NCM, nearest class mean; RF, random forest.

Using the entire ^{1}H NMR spectrum as a fingerprint in conjunction with LDA, Godelmann et al. classified German wines from five regions according to geography, variety and vintage with overall correct classifications of 89% (geographical), 95% (varietal), and 96–97% (vintage) [52]. ^{1}H NMR metabolomic data has also been applied for quantification of a range of metabolites including sugars, amino acids, organic acids, alcohols, and phenolic compounds, which were used for wine discrimination as a function of terroir (encompassing biophysical and cultural factors of the production region) and cultivar [55].

Moreover, Alexandra et al. explored the possibility of combining untargeted ^{1}H NMR analysis with targeted peptide based sensing arrays to classify Pinot noir wines on the basis of characteristic metabolic signatures associated with variations in terroir [56]. Other recent studies have also used ^{1}H NMR as a nontargeted method for authentication. Fan et al. subjected 99 red and 71 white wines from China to NMR analysis, subsequently using segment-wise peak alignment followed by PCA and LDA for separating red and white wine samples as well as different varieties [53]. Similarly, Mascellani et al. used NMR to classify over 900 Czech wines according to type (based on colour and residual sweetness) and variety using a random forest (RF) machine learning algorithm [54]. Correct classification according to wine type was 92% or more for white wine styles (dry and medium dry, medium, sweet) and > 99% for dry red, but the chosen model was unable to provide correct classification for all varieties, with some varieties such as Sauvignon blanc, Pinot Gris, Pinot blanc, and Pálava being below 50% accuracy. Overall, NMR is shown to be an effective technique for authentication with rapid determination of range of metabolites, even if it has become the most expensive spectroscopic approach [33].

In the selection of techniques, it is important to consider the various merits and characteristics of the approaches and to evaluate these according to the question to be addressed. Thus, despite the potential challenges, each of the reviewed methods prevail due to their usability in wine authentication. A summary of the techniques including perceived advantages and disadvantages is presented in Table 6.

Table 6. Summary of spectroscopic techniques applied to wine authentication [33,57].

Technique	Chemical Marker	Advantages	Disadvantages
UV–Vis	Hydroxybenzoic acids, hydroxycinnamic acids, flavan-3-ols, flavonols, and anthocyanin glucosides	Simple analysis, low cost, small volume	Difficulty in identifying specific analytes
IR	Organic acids, alcohols, reducing sugars, and polyphenols	Rapid, simple, qualitative and quantitative analysis	Sensitive to pH and temperature, high interference of water (NIR)
Raman	Organic acids, alcohols, sugars, phenolics	Rapid, small volume, low impact of water	Weak signals, extensive pre-processing requirements
Fluorescence	Phenolics, pigments, vitamins, amino acids	Rapid, sensitive and selective, qualitative and quantitative analysis	Extensive pre-processing requirements, quenching effect
NMR	Phenolics, alcohols, organic acids, amino acids, sugars	Rapid, selective, repeatable and reproducible	Costly equipment, experienced analyst required

3. Application of Chemometrics for Modelling with Spectroscopic Data

Spectroscopic methods rapidly produce an abundance of variables (peak intensities and wavelengths) that need to be dealt with. Therefore, to analyse these high dimensional sets of 'big data', integration with appropriate multivariate statistical analysis methods (i.e., chemometrics) is essential for pattern recognition or modelling (see Figure 3 for an overall approach).

As an exploratory technique that reveals underlying patterns in the data, principal component analysis (PCA) is the most widely applied unsupervised method [33]. It explores the relationship between individual observations and reveals the trends, or groups within the multivariate space [58]. PCA is also applied as a dimension reduction technique that explains the variance of the data matrix in terms of principal components, those being a small number of non-dependent factors containing important information from the original set [59]. Other than differentiating among samples and potentially revealing clustering according to region of origin, for example, data compression with PCA can also be useful prior to other statistical treatments [58]. Notably, PCA is used for two-way array data. With three-way data such as EEMs arising from total fluorescence spectroscopy, PARAFAC can be used instead to decompose and extract the information into different components that

describe the variability of the EEM data more specifically [47]. These aspects are revealed in Figure 3 as early steps in the overall data analysis process.

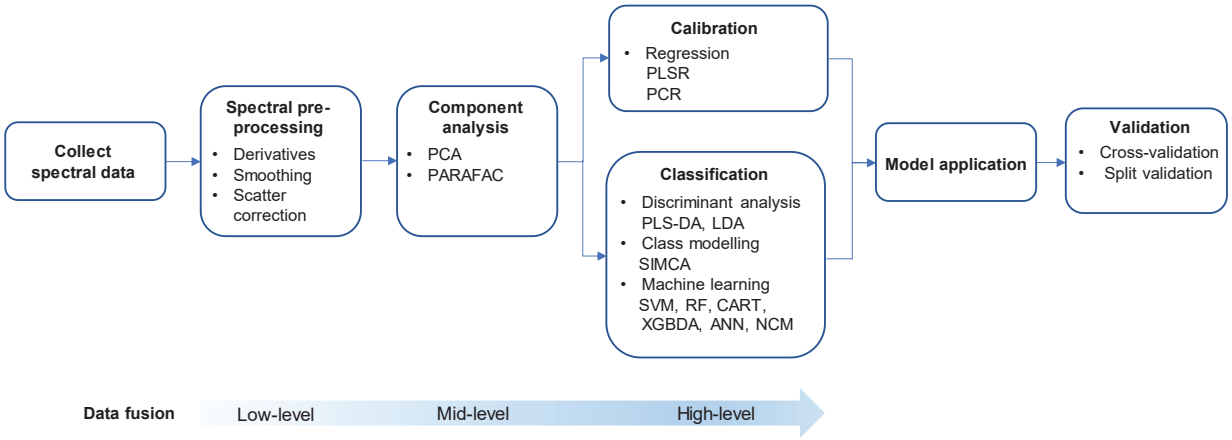


Figure 3. Schematic of the steps involved with chemometric analysis of spectroscopic data. Stages where data fusion can be applied depending on the extent of processing are also presented in the diagram.

For classification purposes, supervised statistical approaches such as discriminant analysis methods are widely applied in authentication of wine (Figure 3). Among spectroscopic studies, PLS-DA and LDA methods have mainly been considered. With LDA (or canonical variate analysis), linear combinations of the original variables (i.e., canonical variates) are estimated that provide maximum separation between classes (groups) while minimising the variance within each class. However, for LDA, the number of training samples needs to be larger than the number of variables, so variable selection by PCA needs to occur with spectroscopic analysis prior to classification with LDA [60]. On the other hand, PLS-DA uses regression to estimate the class of a sample from the variables obtained from a spectral technique, whereby the entire data matrix is regressed on a binary-coded response array and samples are classified according to their predicted values. In their study, Geană et al. showed that LDA works well for classification according to variety with UV–Vis data and PLS-DA improved the classification with FT-IR data [27]. The disadvantage of PLS-DA is that a sample can remain unclassified if it does not belong to any of the pre-defined classes [61].

Another commonly applied supervised technique for classifying wine involves class modelling (Figure 3), and specifically SIMCA, in which similarities among samples belonging to the same class are captured. As explained by Suciu et al., SIMCA is built around PCA and is sensitive enough to identify false outliers to improve the robustness of the model [47]. The advantage of SIMCA over discriminant analyses is that it defines the acceptance area around the target class, which enables delimiting of the target objects from any other objects and classes, and allows assignment of a new sample if it locates in the assigned area of the class [61]. However, due to overlapping of regions, some samples might be classified in one or more classes, and as Rodionova et al. concluded, all classification tasks require the use of an appropriate chemometric approach [61]. That will include the application of new methods, and indeed in more recent years the development of machine learning techniques has shown great potential as they offer advantages in classification compared to conventional methods.

Machine learning started gaining attention in food analysis due to the possibility of performing both linear and non-linear classifications [33]. Among the approaches (Figure 3), SVM has been explored more for wine authentication [62], in conjunction with UV–Vis (Table 1) and NIR (Table 2). SVM is an effective machine learning technique suitable for both classification and regression analysis. It is based on a kernel extension of a binary linear classifier that classifies samples in a hyperplane built according to the features [63].

When a sample set is not balanced over the classes, however, the classification accuracy from SVM may be affected. An alternative machine learning technique involves a decision tree (DT) approach, of which there are several variations, with the most well-known being classification and regression tree (CART). DT methods divide the samples into classes based on the value of certain variables and can be boosted (iterative model) or bagged (independent models including RF), whereby the DT modelling is repeated on random subsets of samples combined into ensembles [64]. Such methods show higher accuracy in classification, are unaffected by outliers or non-linear relationships, and can suitably address class imbalance problems [64]. XGBDA is one such algorithm based on a boosted DT that has recently been applied for the first time (Table 3) to geographical and varietal authentication of red wine using fluorescence spectroscopy [11,48]. Another method for non-linear classification is nearest class mean (NCM), which has rarely been applied in spectroscopic analysis of wine. After dimension reduction of data (from NMR for example, Table 5) with PCA followed by LDA to maximise class separation, NCM can then be used to assign wines to a class with the minimum distance between the respective model class mean and the test-set object [52]. Artificial neural network (ANN) is another option that performs well in classifying samples with non-linear behaviour [65] and has shown acceptable results in variety classification of grapes using FTIR [66]. Although ANN (and CART) has been applied to classification of wine based on anthocyanins [67] or volatiles [68] using chromatographic techniques, there did not appear to be any examples involving spectroscopic data.

Steps of Chemometric Analysis

It is important to appreciate the key stages in any chemometric approach that need to be followed to complete the process (Figure 3).

Apart from the modelling aspects mentioned in the preceding paragraphs, applying a proper spectral pre-processing method depends on the nature of the data set. Noise reduction and baseline offset are common for all spectroscopic techniques and mainly involve smoothing using techniques like the Savitzky-Golay algorithm [69]. For vibrational spectroscopic data, multiplicative scatter correction and standard normal variate methods are utilised for applying corrections to the spectra by comparing signal intensities to a reference signal. Instead for EEM data, correction of Rayleigh masking, Raman scattering, and IFE corrections need to be used. Other than the analytical artefacts, issues can arise with sample variations. For these, it is important to apply pre-processing methods such as normalisation to remove differences due to dilution and for equalising the integral of peaks of the spectra. Different scaling methods, such as autoscaling, and transformations like mean centring are useful in identifying the important variables among others [69].

Data fusion is another practice that can be carried out to enhance the classification of products and predict their properties. After data pre-processing, data fusion (usually involving variables from complementary techniques) can be carried out in different ways. As a relatively simple approach, low-level data fusion (Figure 3) uses measurements directly from different techniques. In contrast, mid-level fusion uses features obtained from the data sources such as PCA scores, which is important when data is diverse in size or scale. In high-level fusion, the results of the different individual models of the data are combined and applied to the classification problem [65].

Another essential aspect of the chemometric application is model validation (Figure 3). After implementation, a classification model's validity has to be verified with a validation sample set, to avoid overfitting of the model and to assess its accuracy. It can be categorised as internal validation when separated into calibration and validation sets, and external validation when independent test sets are used. Cross-validation (CV) is the most commonly applied validation method, consisting of different techniques such as leave-one-out CV, multi-fold (*k*-fold)/Venetian blinds CV, contiguous blocks, and random subsets. There can also be split validation, where the whole data set is divided according to different methods such as random, duplex or Kennard-Stone, [64]. In selecting a suitable validation method,

it is important to consider the number of samples in the set, otherwise validation can lead to inaccurate results in prediction of new samples. For example, if the sample set is not large enough, CV methods would be more suitable over split analysis [70]. Furthermore when considering the sample size, classification methods have been identified as being less susceptible to sample size variation, such as the RF technique [71]. However, it is difficult to suggest which combination of classification/validation approaches would always give significantly better results than any other [72].

Whichever methods are chosen, performance indicators are important parameters to consider in the model validation process. In authentication applications, misclassification and correct classification rates, expressed as a percentage of all samples in a class or an overall average, are most commonly applied to evaluate the model performance in the studies reviewed in this paper. Other than these measures of accuracy, sensitivity and specificity can also be evaluated as performance indicators [48]. Measures of performance specifically for multivariate regression models include coefficient of determination, which represents the goodness of fit of the model based on the training set, and the root mean square errors of calibration and prediction (RMSEC and RMSEP), which are used to understand the predictive capability of the models [64].

4. Future Trends and Directions

Given the international nature of the modern wine trade, reliable methods for assessment of wine authenticity are required to guarantee customer satisfaction of product quality. Potential approaches need to satisfy a number of criteria, foremost of which is having sensitivity to accurately classify non-authentic wines with a high degree of certainty without misclassifying authentic wines as fraudulent. Ideally, a suitable method also needs to be rapid and easily applied, even in a supply chain setting. Therefore, spectroscopic techniques are destined to play a major role due to meeting criteria such as being rapid, user-friendly, and cost-effective.

Among the range of current spectral tools, there have been a number of breakthroughs in the application of spectroscopy for wine analysis. One exciting development is the ability to undertake non-destructive wine measurements through-bottle using various spectroscopic techniques (NIR-Vis, Raman, NMR) which has been successful to a certain extent in identifying oxidation and illegal or hazardous contaminants [73]. Nevertheless, improvements in available techniques or development of new ones to identify chemical markers for geographical, varietal, or vintage authentication is ongoing. NMR provides a powerful platform but is not readily deployable in the production or supply chains, in contrast to things like NIR and UV–Vis. Most recently, great promise has been shown with fluorescence spectroscopy with XGBDA modelling, and indeed, the application of powerful chemometric methods such as machine learning algorithms along with spectroscopic data could be exploited further. Improving the user-friendliness of the statistical techniques is important, however, as that will permit non-specialists to apply them within industry. This could conceivably be solved with the development of cloud-based processing and database management, which could also provide accessibility for authorities for the construction of a robust authenticity database containing rigorous details. Ultimately, integration of innovative technology and modelling approaches will add a new dimension to wine authentication and improve the functionality of the current processes. Importantly, this will give consumers added confidence that the wines they purchase and consume are authentic.

Author Contributions: R.K.R.R.: Conceptualisation, funding acquisition, investigation, visualisation, writing—original draft, writing—review and editing. D.L.C.: Conceptualisation, supervision, writing—review and editing. S.E.P.B.: Conceptualisation, funding acquisition, supervision, writing—review and editing. D.C.: Conceptualisation, writing—review and editing. D.W.J.: Conceptualisation, funding acquisition, project administration, resources, supervision, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: R.K.R.R. was a recipient of an Australian Government Research Training Program Scholarship and a Wine Australia supplementary scholarship (WA Ph1909). The Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au, accessed on 15 June 2021; project number IC170100008) is funded by the Australian Government with additional support from Wine Australia, Waite Research Institute, and industry partners. The University of Adelaide is a member of the Wine Innovation Cluster.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Markoski, M.M.; Garavaglia, J.; Oliveira, A.; Olivaes, J.; Marcadenti, A. Molecular properties of red wine compounds and cardiometabolic benefits. *Nutr. Metab. Insights* **2016**, *9*, 51–57. [CrossRef]
2. Waterhouse, A.L.; Sacks, G.L.; Jeffery, D.W. Introduction. In *Understanding Wine Chemistry*; Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2016.
3. Médina, B.; Salagoïty, M.H.; Guyon, F.; Gaye, J.; Hubert, P.; Guillaume, F. Using new analytical approaches to verify the origin of wine. In *New Analytical Approaches for Verifying the Origin of Food*; Brereton, P., Ed.; Woodhead Publishing: Cambridge, UK, 2013; pp. 149–188.
4. Waterhouse, A.L.; Sacks, G.L.; Jeffery, D.W. Authentication. In *Understanding Wine Chemistry*; Waterhouse, A.L., Sacks, G.L., Jeffery, D.W., Eds.; John Wiley & Sons, Ltd.: Chichester, UK, 2016; pp. 383–391.
5. Popîrdă, A.; Luchian, C.E.; Cotea, V.V.; Colibaba, L.C.; Scutarașu, E.C.; Toader, A.M. A review of representative methods used in wine authentication. *Agriculture* **2021**, *11*, 225. [CrossRef]
6. Arslan, M.; Tahir, H.E.; Zareef, M.; Shi, J.; Rakha, A.; Bilal, M.; Xiaowei, H.; Zhihua, L.; Xiaobo, Z. Recent trends in quality control, discrimination and authentication of alcoholic beverages using nondestructive instrumental techniques. *Trends Food Sci. Technol.* **2021**, *107*, 80–113. [CrossRef]
7. Chapman, J.; Gangadoo, S.; Truong, V.K.; Cozzolino, D. Spectroscopic approaches for rapid beer and wine analysis. *Curr. Opin. Food Sci.* **2019**, *28*, 67–73. [CrossRef]
8. Grijalva-Verdugo, C.; Hernández-Martínez, M.; Meza-Márquez, O.G.; Gallardo-Velázquez, T.; Osorio-Revilla, G. FT-MIR spectroscopy and multivariate analysis for determination of bioactive compounds and antioxidant capacity in Cabernet Sauvignon wines. *CyTA-J. Food* **2018**, *16*, 561–569. [CrossRef]
9. Cayuela, J.A.; Puertas, B.; Cantos-Villar, E. Assessing wine sensory attributes using Vis/NIR. *Eur. Food Res. Technol.* **2017**, *243*, 941–953. [CrossRef]
10. Teixeira dos Santos, C.A.; Pascoa, R.; Sarraguca, M.C.; Porto, P.; Cerdeira, A.L.; Gonzalez-Saiz, J.M.; Pizarro, C.; Lopes, J.A. Merging vibrational spectroscopic data for wine classification according to the geographic origin. *Food Res. Int.* **2017**, *102*, 504–510. [CrossRef]
11. Ranaweera, R.K.R.; Gilmore, A.M.; Capone, D.L.; Bastian, S.E.P.; Jeffery, D.W. Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. *Food Chem.* **2021**, *335*, 127592. [CrossRef] [PubMed]
12. van Eck, N.J.; Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **2010**, *84*, 523–538. [CrossRef] [PubMed]
13. Christoph, N.; Hermann, A.; Wachter, H. 25 years authentication of wine with stable isotope analysis in the European Union—Review and outlook. In *38th World Congress of Vine and Wine*; Jean-Marie, A., Ed.; BIO Web of Conferences; EDP Sciences: Yulis, France, 2015; Volume 5.
14. Ranaweera, R.K.R.; Souza Gonzaga, L.; Capone, D.L.; Bastian, S.E.P.; Jeffery, D.W. 3.33—Authenticity and traceability in the wine industry: From analytical chemistry to consumer perceptions. *Compr. Foodomics* **2021**, *3*, 452–480. [CrossRef]
15. Hofmann, A. Spectroscopic techniques. In *Principles and Techniques of Biochemistry and Molecular Biology*, 7th ed.; Walker, J., Wilson, K., Eds.; Cambridge University Press: Cambridge, UK, 2010; pp. 477–521.
16. Aleixandre-Tudo, J.L.; Buica, A.; Nieuwoudt, H.; Aleixandre, J.L.; du Toit, W. Spectrophotometric analysis of phenolic compounds in grapes and wines. *J. Agric. Food Chem.* **2017**, *65*, 4009–4026. [CrossRef] [PubMed]
17. Yu, J.; Wang, H.; Zhan, J.; Huang, W. Review of recent UV–Vis and infrared spectroscopy researches on wine detection and discrimination. *Appl. Spectrosc. Rev.* **2018**, *53*, 65–86. [CrossRef]
18. Acevedo, F.J.; Jiménez, J.; Maldonado, S.; Domínguez, E.; Narváez, A. Classification of wines produced in specific regions by UV−visible spectroscopy combined with support vector machines. *J. Agric. Food Chem.* **2007**, *55*, 6842–6849. [CrossRef]
19. Azcarate, S.M.; Cantarelli, M.Á.; Pellerano, R.G.; Marchevsky, E.J.; Camiña, J.M. Classification of Argentinean Sauvignon blanc wines by UV spectroscopy and chemometric methods. *J. Food Sci.* **2013**, *78*, C432–C436. [CrossRef]
20. Urbano, M.; Luque de Castro, M.D.; Pérez, P.M.; García-Olmo, J.; Gómez-Nieto, M.A. Ultraviolet–visible spectroscopy and pattern recognition methods for differentiation and classification of wines. *Food Chem.* **2006**, *97*, 166–175. [CrossRef]
21. Philippidis, A.; Poulakis, E.; Kontzedaki, R.; Orfanakis, E.; Symianaki, A.; Zoumi, A.; Velegrakis, M. Application of ultraviolet-visible absorption spectroscopy with machine learning techniques for the classification of Cretan wines. *Foods* **2021**, *10*, 9. [CrossRef]

22. Tan, J.; Li, R.; Jiang, Z.-T.; Zhang, Y.; Hou, Y.-M.; Wang, Y.-R.; Wu, X.; Gong, L. Geographical classification of Chinese Cabernet Sauvignon wines by data fusion of ultraviolet–visible and synchronous fluorescence spectroscopies: The combined use of multiple wavelength differences. *Aust. J. Grape Wine Res.* **2016**, *22*, 358–365. [CrossRef]
23. Cozzolino, D.; Cynkar, W.; Shah, N.; Smith, P.A. Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food Chem.* **2011**, *126*, 673–678. [CrossRef]
24. Cozzolino, D.; Cynkar, W.; Shah, N.; Smith, P. Technical solutions for analysis of grape juice, must, and wine: The role of infrared spectroscopy and chemometrics. *Anal. Bioanal. Chem.* **2011**, *401*, 1475–1484. [CrossRef]
25. Aleixandre-Tudó, J.L.; Castelló-Cogollos, L.; Aleixandre, J.L.; Aleixandre-Benavent, R. Bibliometric insights into the spectroscopy research field: A food science and technology case study. *Appl. Spectrosc. Rev.* **2020**, *55*, 873–906. [CrossRef]
26. Bevin, C.J.; Dambergs, R.G.; Fergusson, A.J.; Cozzolino, D. Varietal discrimination of Australian wines by means of mid-infrared spectroscopy and multivariate analysis. *Anal. Chim. Acta* **2008**, *621*, 19–23. [CrossRef]
27. Geană, E.-I.; Ciucure, C.T.; Apetrei, C.; Artem, V. Application of spectroscopic UV–Vis and FT-IR screening techniques coupled with multivariate statistical analysis for red wine authentication: Varietal and vintage year discrimination. *Molecules* **2019**, *24*, 4166. [CrossRef]
28. Rodriguez-Saona, L.E.; Allendorf, M.E. Use of FTIR for rapid authentication and detection of adulteration of food. *Annu. Rev. Food Sci. Technol.* **2011**, *2*, 467–483. [CrossRef]
29. Cozzolino, D. The role of visible and infrared spectroscopy combined with chemometrics to measure phenolic compounds in grape and wine samples. *Molecules* **2015**, *20*, 726–737. [CrossRef]
30. Martelo-Vidal, M.J.; Domínguez-Agis, F.; Vázquez, M. Ultraviolet/visible/near-infrared spectral analysis and chemometric tools for the discrimination of wines between subzones inside a controlled designation of origin: A case study of Rías Baixas. *Aust. J. Grape Wine Res.* **2013**, *19*, 62–67. [CrossRef]
31. Hu, X.-Z.; Liu, S.-Q.; Li, X.-H.; Wang, C.-X.; Ni, X.-L.; Liu, X.; Wang, Y.; Liu, Y.; Xu, C.-H. Geographical origin traceability of Cabernet Sauvignon wines based on Infrared fingerprint technology combined with chemometrics. *Sci. Rep.* **2019**, *9*, 8256. [CrossRef]
32. Patz, C.D.; Blieke, A.; Ristow, R.; Dietrich, H. Application of FT-MIR spectrometry in wine analysis. *Anal. Chim. Acta* **2004**, *513*, 81–89. [CrossRef]
33. Ríos-Reina, R.; Camiña, J.M.; Callejón, R.M.; Azcarate, S.M. Spectralprint techniques for wine and vinegar characterization, authentication and quality control: Advances and projections. *TrAC Trends Anal. Chem.* **2021**, *134*, 116121. [CrossRef]
34. Magdas, D.A.; Guyon, F.; Feher, I.; Pinzaru, S.C. Wine discrimination based on chemometric analysis of untargeted markers using FT-Raman spectroscopy. *Food Control.* **2018**, *85*, 385–391. [CrossRef]
35. Magdas, D.A.; Cinta Pinzaru, S.; Guyon, F.; Feher, I.; Cozar, B.I. Application of SERS technique in white wines discrimination. *Food Control.* **2018**, *92*, 30–36. [CrossRef]
36. Deneva, V.; Bakardzhiyski, I.; Bambalov, K.; Antonova, D.; Tsobanova, D.; Bambalov, V.; Cozzolino, D.; Antonov, L. Using Raman spectroscopy as a fast tool to classify and analyze Bulgarian wines—A feasibility study. *Molecules* **2020**, *25*, 170. [CrossRef]
37. Magdas, D.A.; Cozar, B.I.; Feher, I.; Guyon, F.; Dehelean, A.; Cinta Pinzaru, S. Testing the limits of FT-Raman spectroscopy for wine authentication: Cultivar, geographical origin, vintage and terroir effect influence. *Sci. Rep.* **2019**, *9*, 19954. [CrossRef]
38. Zanuttin, F.; Gurian, E.; Ignat, I.; Fornasaro, S.; Calabretti, A.; Bigot, G.; Bonifacio, A. Characterization of white wines from north-eastern Italy with surface-enhanced Raman spectroscopy. *Talanta* **2019**, *203*, 99–105. [CrossRef]
39. de Almeida, M.P.; Leopold, N.; Franco, R.; Pereira, E. Expedite SERS fingerprinting of Portuguese white wines using plasmonic silver nanostars. *Front. Chem.* **2019**, *7*. [CrossRef]
40. Xu, Y.; Zhong, P.; Jiang, A.; Shen, X.; Li, X.; Xu, Z.; Shen, Y.; Sun, Y.; Lei, H. Raman spectroscopy coupled with chemometrics for food authentication: A. review. *TrAC Trends Anal. Chem.* **2020**, *131*, 116017. [CrossRef]
41. Sádecká, J.; Tothova, J. Fluorescence spectroscopy and chemometrics in the food classification: A review. *Czech. J. Food Sci.* **2007**, *25*, 159–174. [CrossRef]
42. Airado-Rodríguez, D.; Durán-Merás, I.; Galeano-Díaz, T.; Wold, J.P. Front-face fluorescence spectroscopy: A new tool for control in the wine industry. *J. Food Compos. Anal.* **2011**, *24*, 257–264. [CrossRef]
43. Airado-Rodríguez, D.; Galeano-Díaz, T.; Durán-Merás, I.; Wold, J.P. Usefulness of fluorescence excitation−emission matrices in combination with PARAFAC, as fingerprints of red wines. *J. Agric. Food Chem.* **2009**, *57*, 1711–1720. [CrossRef]
44. Coelho, C.; Aron, A.; Roullier-Gall, C.; Gonsior, M.; Schmitt-Kopplin, P.; Gougeon, R.D. Fluorescence fingerprinting of bottled white wines can reveal memories related to sulfur dioxide treatments of the must. *Anal. Chem.* **2015**, *87*, 8132–8137. [CrossRef]
45. Dankowska, A. Advances in fluorescence emission spectroscopy for food authenticity testing. In *Advances in Food Authenticity Testing*; Woodhead Publishing: Cambridge, UK, 2016; pp. 117–145.
46. Sádecká, J.; Jakubíková, M. Varietal classification of white wines by fluorescence spectroscopy. *J. Food Sci. Technol.* **2020**, *57*, 2545–2553. [CrossRef] [PubMed]
47. Suciu, R.-C.; Zarbo, L.; Guyon, F.; Magdas, D.A. Application of fluorescence spectroscopy using classical right angle technique in white wines classification. *Sci. Rep.* **2019**, *9*, 1–10. [CrossRef]
48. Ranaweera, R.K.R.; Gilmore, A.M.; Capone, D.L.; Bastian, S.E.P.; Jeffery, D.W. Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. *Food Chem.* **2021**, 130149. [CrossRef] [PubMed]

49. Martin, G.J.; Martin, M.L.; Mabon, F.; Michon, M.J. A new method for the identification of the origin of ethanols in grain and fruit spirits: High-field quantitative deuterium nuclear magnetic resonance at the natural abundance level. *J. Agric. Food Chem.* **1983**, *31*, 311–315. [CrossRef]
50. Solovyev, P.A.; Fauhl-Hassek, C.; Riedl, J.; Esslinger, S.; Bontempo, L.; Camin, F. NMR spectroscopy in wine authentication: An official control perspective. *Compr. Rev. Food Sci. Food Saf.* **2021**, *20*, 2040–2062. [CrossRef] [PubMed]
51. Viskić, M.; Bandić, L.M.; Korenika, A.-M.J.; Jeromel, A. NMR in the service of wine differentiation. *Foods* **2021**, *10*, 120. [CrossRef]
52. Godelmann, R.; Fang, F.; Humpfer, E.; Schütz, B.; Bansbach, M.; Schäfer, H.; Spraul, M. Targeted and nontargeted wine analysis by ^{1}H NMR spectroscopy combined with multivariate statistical analysis. Differentiation of important parameters: Grape variety, geographical origin, year of vintage. *J. Agric. Food Chem.* **2013**, *61*, 5610–5619. [CrossRef]
53. Fan, S.; Zhong, Q.; Fauhl-Hassek, C.; Pfister, M.K.H.; Horn, B.; Huang, Z.B. Classification of Chinese wine varieties using ^{1}H NMR spectroscopy combined with multivariate statistical analysis. *Food Control* **2018**, *88*, 113–122. [CrossRef]
54. Mascellani, A.; Hoca, G.; Babisz, M.; Krska, P.; Kloucek, P.; Havlik, J. ^{1}H NMR chemometric models for classification of Czech wine type and variety. *Food Chem.* **2021**, *339*, 127852. [CrossRef] [PubMed]
55. Gougeon, L.; Da Costa, G.; Le Mao, I.; Ma, W.; Teissedre, P.-L.; Guyon, F.; Richard, T. Wine analysis and authenticity using ^{1}H-NMR metabolomics data: Application to Chinese wines. *Food Anal. Methods* **2018**, *11*, 3425–3434. [CrossRef]
56. Crook, A.A.; Zamora-Olivares, D.; Bhinderwala, F.; Woods, J.; Winkler, M.; Rivera, S.; Shannon, C.E.; Wagner, H.R.; Zhuang, D.L.; Lynch, J.E.; et al. Combination of two analytical techniques improves wine classification by vineyard, region, and vintage. *Food Chem.* **2021**, *354*, 129531. [CrossRef]
57. Sáenz Navajas, M.-P.; Jeffery, D.W. Perspectives on wines of provenance: Sensory typicality, quality, and authenticity. *ACS Food Sci. Technol.* **2021**, *1*, 986–992. [CrossRef]
58. Rodionova, O.; Kucheryavskiy, S.; Pomerantsev, A. Efficient tools for principal component analysis of complex data— A tutorial. *Chemom. Intell. Lab. Syst.* **2021**, *213*, 104304. [CrossRef]
59. Cozzolino, D.; Power, A.; Chapman, J. Interpreting and reporting principal component analysis in food science analysis and beyond. *Food Anal. Methods* **2019**, *12*, 2469–2473. [CrossRef]
60. Brito, A.L.B.; Brito, L.R.; Honorato, F.A.; Pontes, M.J.C.; Pontes, L.F.B.L. Classification of cereal bars using near infrared spectroscopy and linear discriminant analysis. *Food Res. Int.* **2013**, *51*, 924–928. [CrossRef]
61. Rodionova, O.Y.; Titova, A.V.; Pomerantsev, A.L. Discriminant analysis is an inappropriate method of authentication. *TrAC Trends Anal. Chem.* **2016**, *78*, 17–22. [CrossRef]
62. Ríos-Reina, R.; Elcoroaristizabal, S.; Ocaña-González, J.A.; García-González, D.L.; Amigo, J.M.; Callejón, R.M. Characterization and authentication of Spanish PDO wine vinegars using multidimensional fluorescence and chemometrics. *Food Chem.* **2017**, *230*, 108–116. [CrossRef]
63. Xu, Y.; Zomer, S.; Brereton, R.G. Support vector machines: A recent method for classification in chemometrics. *Crit. Rev. Anal. Chem.* **2006**, *36*, 177–188. [CrossRef]
64. Jiménez-Carvelo, A.M.; González-Casado, A.; Bagur-González, M.G.; Cuadros-Rodríguez, L. Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity—A review. *Food Res. Int.* **2019**, *122*, 25–39. [CrossRef]
65. Borràs, E.; Ferré, J.; Boqué, R.; Mestres, M.; Aceña, L.; Busto, O. Data fusion methodologies for food and beverage authentication and quality assessment—A review. *Anal. Chim. Acta* **2015**, *891*, 1–14. [CrossRef]
66. Murru, C.; Chimeno-Trinchet, C.; Díaz-García, M.E.; Badía-Laíño, R.; Fernández-González, A. Artificial Neural Network and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy to identify the chemical variables related to ripeness and variety classification of grapes for protected. Designation of origin wine production. *Comput. Electron. Agric.* **2019**, *164*, 104922. [CrossRef]
67. Cosme, F.; Milheiro, J.; Pires, J.; Guerra-Gomes, F.I.; Filipe-Ribeiro, L.; Nunes, F.M. Authentication of Douro DO monovarietal red wines based on anthocyanin profile: Comparison of partial least squares—Discriminant analysis, decision trees and artificial neural networks. *Food Control.* **2021**, *125*, 107979. [CrossRef]
68. Kruzlicova, D.; Mocak, J.; Balla, B.; Petka, J.; Farkova, M.; Havel, J. Classification of Slovak white wines using artificial neural networks and discriminant techniques. *Food Chem.* **2009**, *112*, 1046–1052. [CrossRef]
69. Engel, J.; Gerretzen, J.; Szymańska, E.; Jansen, J.J.; Downey, G.; Blanchet, L.; Buydens, L.M.C. Breaking with trends in pre-processing? *TrAC Trends Anal. Chem.* **2013**, *50*, 96–106. [CrossRef]
70. Callao, M.P.; Ruisanchez, I. An overview of multivariate qualitative methods for food fraud detection. *Food Control.* **2018**, *86*, 283–293. [CrossRef]
71. Moghaddam, D.D.; Rahmati, O.; Panahi, M.; Tiefenbacher, J.; Darabi, H.; Haghizadeh, A.; Haghighi, A.T.; Nalivan, O.A.; Tien Bui, D. The effect of sample size on different machine learning models for groundwater potential mapping in mountain bedrock aquifers. *CATENA* **2020**, *187*, 104421. [CrossRef]
72. Xu, Y.; Goodacre, R. On splitting training and validation set: A comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. *J. Anal. Test.* **2018**, *2*, 249–262. [CrossRef]
73. Augustine, M.P.; Harley, S.J.; Lim, V.; Stucky, P. An Authentication Device for Full Intact Wine Bottles. U.S. Patent 9,488,599, 8 November 2016.

Research Questions

As a luxury product, the economic value of wine is linked with culture and lifestyle, and consumers specifically demand reliable information about product origins. Considering the cost associated with wine fraud on a global scale, techniques that can authenticate wine based on provenance, variety, and vintage are necessary for brand protection as well as to enhance consumer confidence. Wine authentication has been a challenge for analytical chemists over the past few decades and despite the major advances, there is still a need for a robust technique that could be applied in an industrial setting rather than a specialised laboratory. On the other hand, understanding the influence of terroir and authenticating markers relevant to provenance will improve the value of regional wine. With a particular focus on Australian wine, understanding the unique characteristics inherent in wine regions and varieties by applying cutting-edge technology will help guarantee the provenance indicated on the label. In terms of practical application, it is desirable to consider a method that is rapid, accessible for in situ analyses, simple to operate, relatively low cost, and has high sensitivity and specificity. Therefore, upon reviewing the literature as presented in the preceding chapters, the following research questions were addressed by this thesis:

1. **Could fluorescence spectroscopy be used in the differentiation of Australian wine according to geographical origin, variety, and vintage?**
   Fluorescence spectroscopy is known as a simple, non-destructive, non-invasive and relatively inexpensive spectroscopic technique. Despite the potential benefits, the application of total fluorescence spectroscopy in wine authentication has been scarcely explored. In combination with the advancement of instrument capabilities, the opportunity arose for the development of a fast, reliable measurement approach and elaboration of fluorescence molecular fingerprints for wine authentication purposes.

2. **What chemometric methods would be applicable for developing classification models for wine authentication and regression models to quantify wine chemical parameters?**

   As with other spectral approaches, chemometrics plays a vital role in modelling fluorescence data. Unsupervised methods are useful for revealing underlying patterns in the data as well as for dimension reduction. For classification, supervised statistical approaches are essential, such as discriminant analysis methods. Machine learning techniques are gaining in popularity and have effectively been applied in discrimination studies. It was necessary to identify suitable methods to develop classification models by considering factors such as sample size, number of classes, and linearity of data. Spectroscopy can also be used as a method for predicting chemical parameters such as pH, TA, alcohol and the concentration of phenolic compounds. The possibility of applying fluorescence data for modelling chemical parameters can be evaluated with different regression methods.

3. **Can the molecular fingerprint embedded in fluorescence spectra be applied as evidence to understand the variation of terroir and explore underlying chemical markers?**

   Having tangible evidence of the influence of geographical origin on fine wine composition as a function of its terroir can provide fundamental scientific understanding. Studies based on sensory analysis, climatic and topographic indices, soil properties, and grape and wine chemistry have been undertaken to assess region influences at various scales, with some being investigated to understand subregional terroirs in South Australian wine regions. However, the possibility of demonstrating variations ascribed to intraregional differences in terroir through chemical analysis still remained unclear. Additionally, understanding chemical markers or compositional variables that drive the differences is important for the optimisation of regional expression of wine. The capability of fluorescence spectroscopy to detect intraregional variation needed to be determined.

4. **Can spectrofluorometric analysis be used to trace the origin of wine throughout the winemaking process and identify the blending percentage of varietal wines?**

   When introducing a method for validating the origin of wine and detecting wine fraud, being able to trace a wine's origin through production in conjunction with identifying small additions of other wine in a blend would be beneficial. To date, Australia has not developed a scientific approach to deal with this challenge in relation to the Label Integrity Program, which stipulates a "wine must contain a minimum of 85 % of grapes from a declared variety and GI". Experiments were required to examine different stages of the winemaking process and wine blended with known percentages of a different variety, to evaluate the ability of fluorescence spectroscopy to address these challenges.

Summary of Research Aims

The overall aim of this project was to implement a reliable analytical method and develop robust models for the classification of Australian red wine according to region of origin. In addition, gaining an understanding of marker compounds encompassed within the chemical composition of wine was considered. The project will be beneficial for investigating chemical signatures in relation to regional expression in wine and developing a procedure that can offer a measure of protection against wine fraud in the global market.

The aims will be addressed through the following objectives:

1. Investigate spectrofluorometric analysis based on an absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) approach and undertake classification modelling for geographical authentication of Australian Cabernet Sauvignon wines from different regions and from Bordeaux as an international benchmark, and compare the effectiveness of the approach with element composition determined by inductively coupled plasma mass spectrometry (ICP-MS) as a reference method.

2. Verify the utility of the spectroscopic methodology for authentication with a wide range of Australian red wines from different regions and varieties by applying appropriate chemometric methods and machine learning techniques to classify and validate models, further identify the chemical markers in relation to wine provenance, and predict the concentration of chemical constituents using spectral data.

3. Explore the possibility of relating the molecular fingerprint of wine with the variations of terroir at a subregional level within a South Australian wine region to improve the understanding of the impact of terroir on intrinsic wine properties, and identify the possibility of discrimination of samples according to vintage using suitable chemometric methods.

4. Determine the feasibility of tracing the original fingerprint of wine throughout the winemaking process and identify the blending percentage of varietal wines in relation to Australia's Label Integrity Program, thereby helping to establish the provenance of wine through chemical traceability.

Chapter 3

Publication

Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling.

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. Food Chemistry, 335, 127592.

Statement of Authorship

Title of Paper	Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling
Publication Status	☑ Published ☐ Accepted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style ☐ Submitted for Publication
Publication Details	Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. Food Chemistry, 335, 127592. https://doi.org/https://doi.org/10.1016/j.foodchem.2020.127592.

Principal Author

Name of Principal Author (Candidate)	Ranaweera K. R. Ranaweera		
Contribution to the Paper	Contributed to the experimental design. Performed analysis of basic chemical parameters, ICP-MS and A-TEEM analysis for the wine samples. Carried out data collection, processing, statistical analysis and interpretation of data for the manuscript. Drafted the manuscript, and reviewed and edited the final draft for publication.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	12/02/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Adam M. Gilmore		
Contribution to the Paper	Contributed to design of methodology for wine analysis. Applied statistical techniques in formal analysis of the study data. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	2021/08/03

Name of Co-Author	Dimitra L. Capone		
Contribution to the Paper	Contributed to the planning and design of experiments. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	20/07/2021

Name of Co-Author	Susan E.P. Bastian		
Contribution to the Paper	Contributed to the planning and design of experiments. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	14.2.2021

Name of Co-Author	David W. Jeffery		
Contribution to the Paper	Original conceptualisation of the study. Acquired funding and provided resources for the project leading to this publication. Supervised the project and undertook project administration throughout the study. Reviewed, edited and revised the manuscript for publication. Acted as the corresponding author at all stages.		
Signature		Date	14/02/2021

Food Chemistry 335 (2021) 127592

Contents lists available at ScienceDirect

Food Chemistry

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

Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling

Ranaweera K.R. Ranaweera[a], Adam M. Gilmore[b], Dimitra L. Capone[a,c], Susan E.P. Bastian[a,c], David W. Jeffery[a,c,*]

[a] Department of Wine and Food Science, and Waite Research Institute, The University of Adelaide (UA), PMB 1, Glen Osmond, South Australia 5064, Australia
[b] HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, United States
[c] Australian Research Council Training Centre for Innovative Wine Production, UA, PMB 1, Glen Osmond, South Australia 5064, Australia

ARTICLE INFO

Keywords:
Fluorescence spectroscopy
Phenolics
Colour
Chemometrics
Excitation-emission matrix
Authenticity

ABSTRACT

With the increased risk of wine fraud, a rapid and simple method for wine authentication has become a necessity for the global wine industry. The use of fluorescence data from an absorbance and transmission excitation-emission matrix (A-TEEM) technique for discrimination of wines according to geographical origin was investigated in comparison to inductively coupled plasma-mass spectrometry (ICP-MS). The two approaches were applied to commercial Cabernet Sauvignon wines from vintage 2015 originating from three wine regions of Australia, along with Bordeaux, France. Extreme gradient boosting discriminant analysis (XGBDA) was examined among other multivariate algorithms for classification of wines. Models were cross-validated and performance was described in terms of sensitivity, specificity, and accuracy. XGBDA classification afforded 100% correct class assignment for all tested regions using the EEM of each sample, and overall 97.7% for ICP-MS. The novel combination of A-TEEM and XGBDA was found to have great potential for accurate authentication of wines.

1. Introduction

Geographical origin is associated directly with the quality and status of agricultural products, including wine. Consequently, consumers are specifically interested in the provenance of the products they consume, due to globalisation and an increase in fraudulent practices in the food industry (Danezis et al., 2016). As a way of identifying products that originate from a particular region, geographical indication (GI) and appellation systems have been introduced for control and protection. These systems recognise the importance of environmental and human factors, associated with vineyards and wineries in the case of wine, which influence the product characteristics.

As a luxury product, wine is vulnerable to counterfeiting. Authentication according to the GI as claimed on the label is therefore vital in addressing the issue of wine fraud, which had an estimated economic cost of several hundred million dollars to Australia in 2017 (McLeod, 2017). Moreover, verification of GI is particularly important because certain grape cultivars can be strongly associated with specific wine regions, with both of these factors being significant drivers of wine quality (Jackson, 2009) due to their influences on grape (and wine) composition.

Wine is compositionally diverse, with chemical components being derived from the grape berry, and yeast and bacterial metabolism, along with the ageing process as a result of oak wood contact or oxidative and acid-catalysed chemical reactions upon storage (Waterhouse, Sacks, & Jeffery, 2016). Specific chemical measures based on compound categories (i.e., volatiles, non-volatiles, elements, isotopes) are important to the authentication of wine, and chemical

Abbreviations: A-TEEM, absorbance-transmission and fluorescence excitation emission matrix; CV, cross-validation; DA, discriminant analysis; EEM, excitation-emission matrix; GI, geographical indication; ICP-MS, inductively coupled plasma-mass spectrometry; IFE, inner filter effects; IRMS, isotope ratio mass spectrometry; MIR, mid-infrared; NIR, near-infrared; PARAFAC, parallel factor analysis; PCA, principal component analysis; PDO, Protected Designation of Origin; SIMCA, soft independent modelling of class analogy; SVM, support vector machine; TA, titratable acidity; TAn, total anthocyanins; TP, total phenolics; XGB, extreme gradient boosting

* Corresponding author.
E-mail address: david.jeffery@adelaide.edu.au (D.W. Jeffery).

https://doi.org/10.1016/j.foodchem.2020.127592
Received 30 April 2020; Received in revised form 13 July 2020; Accepted 13 July 2020
Available online 17 July 2020

markers of geographical origin and variety are often a primary focus (Versari, Laurie, Ricci, Laghi, & Parpinello, 2014). As such, wine authentication has been explored using elemental composition by inductively coupled plasma-mass spectrometry (ICP-MS) (Coetzee, van Jaarsveld, & Vanhaecke, 2014; Martin, Watling, & Lee, 2012), stable isotope ratios by isotope ratio mass spectrometry (IRMS) (Kokkinofta, et al., 2017), amino acid profile (Herbert, Barros, & Alves, 2000), and grape and wine volatile compounds (Villano et al., 2017) and phenolic compounds by MS coupled with liquid or gas chromatography.

Elemental profiling has been identified as one of the most reliable methods for assessing geographical origin due to the relationship between the soil's elemental composition with that of the wine via the grapes produced in that soil. However, elements such as Cu and Zn from fungicides or Na and Ca from the vinification process (e.g., addition of bentonite) might also influence the wine's elemental profile (Jurado, Alcázar, Palacios-Morillo, & de Pablos, 2012; Waterhouse, Sacks, & Jeffery, 2016). In conjunction with chemometrics, studies have successfully used mineral elements in wine as an authentication marker in combination with stable isotopes such as O and C (Kokkinofta, et al., 2017) and heavy metal isotopes like Pb and Sr (Bora, Donici, Teodor, Bunea, Popescu, & Bunea, 2017). Factors such as climatic conditions, geography, and viticultural practices also affect the biosynthesis and accumulation of phenolic compounds in grape berries (and phenolic composition in wine) (Obreque-Slier et al., 2010). Thus, phenolic profile has been recognised as an indicator for differentiation of wine according to origin (Geana et al., 2014; Jaitz et al., 2010), although a requirement for high-resolution techniques and long analysis times would appear to have limited the application of such methods for routine authentication.

More recently, greater attention has been given to the applications of spectroscopic methods for wine classification, due to these being rapid, cost-effective, and simplified methods that assure a high level of robustness and precision (Lohumi, Lee, Lee, & Cho, 2015). A combination of UV, visible (vis), near-infrared (NIR) and mid-infrared (MIR) spectroscopy with multivariate (chemometric) data analysis was explored as a tool to classify commercial Sauvignon Blanc wines from Australia and New Zealand according to geographical origin (Cozzolino, Cynkar, Shah, & Smith, 2011) as well as for Shiraz wines from different Australian regions (Riovanto, Cynkar, Berzaghi, & Cozzolino, 2011). In their study of Sauvignon Blanc wines, Cozzolino et al. (2011) achieved 93% correct classification of the Australian wines using NIR and MIR in combination. As an alternative, proton nuclear magnetic resonance (^{1}H NMR) spectroscopy has also been widely used in combination with multivariate statistical analysis for the classification of wine samples from different regions (Amargianitaki & Spyros, 2017). However, NMR has the disadvantage of only being capable of detecting metabolites that are usually present in higher concentrations (Pinu, 2018).

In contrast to the more traditional approaches, exploration of the potential of fluorescence spectroscopy as a means of wine authentication has been limited over the years (Azcarate et al., 2015; Dufour, Letort, Laguet, Lebecque, & Serra, 2006; Suciu, Zarbo, Guyon, & Magdas, 2019; Yin, Li, Ding, & Wang, 2009). Fluorescence spectroscopy has the desired characteristics of being a rapid and highly sensitive method that is cost-effective. It can reveal specific fluorophore-containing compounds such as phenolics, vitamins, and aromatic amino acids according to particular excitation and emission wavelengths (Airado-Rodríguez, Durán-Merás, Galeano-Díaz, & Wold, 2011). Furthermore, this method enables the "fingerprint" of the samples to be obtained according to the fluorophoric molecular and macromolecular compounds by recording a three-dimensional excitation-emission matrix (EEM) over multiple excitation and emission wavelengths (Coelho, Aron, Roullier-Gall, Gonsior, Schmitt-Kopplin, & Gougeon, 2015). Such an approach coupled with multivariate data analysis techniques has been exploited in discriminating different products according to their geographical origin, such as patchouli oil (Al Riza, Widodo, Purwanto, & Kondo, 2019), wine vinegar (Ríos-Reina et al., 2017) and white wine (Suciu, Zarbo, Guyon, & Magdas, 2019).

EEM can be recorded simultaneously with classic UV–vis absorbance data using the A-TEEM technique developed by Horiba Scientific, which has typically been used for the analysis of organic matter in water but can also be applied to phenolics in wine, among other matrices (Quatela et al., 2018). In terms of data analysis, generally, a variable reduction method such as principal component analysis (PCA) is carried out to increase the effective visualisation of data. These preprocessed data can be explored by a decomposition model such as parallel factor analysis (PARAFAC) to describe the variability of EEMs, and subsequently utilised in differentiation of the samples according to origin with different discrimination models. For example, Azcarate et al. (2015) discriminated Argentinean white wines according to variety, beginning with PCA and PARAFAC, and then using different algorithms such as soft independent modelling of class analogy (SIMCA) with a classification accuracy of 50%, and multiple linear regression methods. The latter included discriminant analysis using multi-way partial least squares, unfolded partial least squares, and successive projection algorithm, affording 65%, 76% and 80% correct classification, respectively. In contrast, machine learning techniques have been employed more effectively in food authentication. For example, in the authentication of Protected Designation of Origin (PDO) wine vinegar, Ríos-Reina et al. (2017) applied a support vector machine (SVM) approach and obtained a correct classification of 92%. Among the machine learning tools, extreme gradient boosting (XGB) has been identified for its fundamental ability to analyse complex descriptors that are heterogeneous or imbalanced in class distribution, but it has rarely been applied in biochemical prediction (Babajide Mustapha & Saeed, 2016).

Based on the hypothesis that fluorescence spectroscopy in conjunction with chemometrics could be used to build robust authentication models for wine, the utility of A-TEEM spectroscopy was investigated as a tool for regional authentication of Australian Cabernet Sauvignon wines along with a comparison of wines from Bordeaux, France. Different chemometric algorithms were assessed for modelling of fluorescence data for geographical origin classification in comparison to using elemental profile obtained from ICP-MS of the same set of wine samples as a reference method. The effectiveness of the cross-validated modelling techniques was compared by considering the specificity, sensitivity and accuracy of the predictions.

2. Material and methods

2.1. Chemicals and solutions

The following chemicals and consumables were obtained from commercial suppliers: nitric acid (HNO_3) trace metal grade 70% (Thermo Fisher Scientific, Scoresby, VIC, Australia); absolute ethanol, gradient HPLC grade, and hydrochloric acid (HCl), 37% analytical grade (Chem-Supply, Port Adelaide, SA, Australia); ICP multi-component standards and ICP single component standard of Indium (100 μg mL^{-1}) in 2% HNO_3 (Choice Analytical, Thornleigh, NSW, Australia). Water used was purified through a Milli-Q purification system (Millipore, North Ryde, NSW, Australia).

2.2. Wine samples

A total of 86 samples (Coonawarra (C) = 36, Margaret River (M) = 20, Yarra Valley (Y) = 20, and Bordeaux (B) = 10) of commercially available Cabernet Sauvignon wine were analysed in this work. All wine samples were from the 2015 vintage and GI certified as single regions. Prior to analysis, wines were subsampled from the bottles into 4 mL glass vials with Teflon lined caps and stored in a refrigerator at 4 °C.

2.3. Analytical procedure for basic chemical measures

Wine pH and titratable acidity (TA) measurements were obtained with a Mettler Toledo T50 autotitrator, and alcohol content (percentage by volume) was measured with an Anton Paar Alcolyser. Colour measurements, total phenolics (TP), total anthocyanins (TAn), and hue were calculated from a modified Somers assay performed with a SpectraMax M2 microplate reader in triplicate as previously described (Mercurio, Dambergs, Herderich, & Smith, 2007).

2.4. Analytical procedure for ICP-MS

HNO_3 and ethanol were used in dilutions and for preparing calibration standards. All glassware was soaked in 2% HNO_3 overnight, rinsed thoroughly with Milli-Q water, and oven-dried before use. Wine samples were diluted 1:10 with 2% HNO_3 and Indium solution (100 μg mL^{-1} in 2% HNO_3) was added as the internal standard. The method blank was prepared with ethanol and 2% HNO_3 with the same concentration of Indium as present in the wine samples. Calibration standards were prepared in a series of 5, 10, 20, 50, 100, 200, 400, 500, and 1000 μg L^{-1} with multi-element ICP standard stock solutions using ethanol and 2% HNO_3 solution as the diluent.

An Agilent 8900 triple quadrupole ICP-MS was used for the analysis with helium and oxygen as collision and reaction gases for the effective removal of spectral interferences to obtain the maximum concentration of the elements in the wine samples in reference to the background-equivalent concentration (Grindlay, Mora, de Loos-Vollebregt, & Vanhaecke, 2014). Possible matrix effects were checked by running 100 and 200 μg L^{-1} calibration standards in between every ten samples as a quality control step. Method blanks were also analysed between every ten samples. A total of 65 elements were measured to identify possible indicators related to geographical origin based on previous studies (Coetzee, van Jaarsveld, & Vanhaecke, 2014; Martin, Watling, & Lee, 2012). After obtaining the concentrations of all elements, signal drift correction was applied according to the calibration standards.

2.5. Analytical procedure for A-TEEM

The wine samples for A-TEEM analysis were first filtered with a 0.45 μm nylon filter and then diluted 1:200 with a solvent comprising 50% ethanol adjusted with HCl to pH 2, to yield a linear Beer-Lambert concentration relationship (Gilmore, 2014). The samples were analysed after 2 min with stirring in a type 1FL (1 cm path length) Macro Fluorescence cuvette (FireflySci, Inc. Staten Island, NY, USA) within the excitation wavelength range of 240–800 nm with a 5 nm increment under medium gain and 0.1 s integration time, and emission wavelength range of 242–824 nm with a 4.66 nm increment as set by the instrument. Absorbance spectra, hue, intensity CIE L*a*b, CIE 1931, and EEMs were recorded for each sample prepared in duplicate using a Horiba Scientific Aqualog spectrophotometer (version 4.2, Quark Photonics, Adelaide, SA, Australia). Aqualog integrates its data acquisition functions directly with the Origin software (version 8.6, OriginLab Corporation, Massachusetts, USA). All EEMs were normalised according to the water Raman scattering units for the specified emission conditions and the data were corrected for the influence of inner filter effects (IFE) and Rayleigh masking prior to statistical analysis (Gilmore, Akaji, & Csatorday, 2017).

2.6. Statistical analyses

XLSTAT was used for one-way analysis of variance (ANOVA) on the basic chemical measures and multi-element data according to region (version 2019.03.02, Addinsoft, Boston, USA). As an unsupervised dimensionality reduction technique, PCA was carried out using a correlation matrix to account for differences in the concentration ranges of elements, and from PCA results after varimax rotation, stepwise (forward) discriminant analysis (DA) with cross-validation was conducted with $\alpha = 0.05$ to discriminate the wine samples according to geographic origin, using XLSTAT in both cases. EEM data analysis was executed by support vector machine discriminant analysis (SVMDA) and pre-processed with different options including mean centring, autoscaling, and class centroid centring, and compressed by PCA or PLS regression to produce latent variables, with components selected that explained the original data with the highest correct classification percentage. As an effective machine learning technique used for building predictive tree-based models, extreme gradient boosting discriminant analysis (XGBDA) was applied for both elements and EEM data sets. After unfolding the EEM data using transform unfold multiway (mode 1) tool in Solo software (version 8.7.1, Eigenvector Research, Inc., Manson, WA, USA), model compression was undertaken with PLS regression, using a maximum of 10 latent variables, with mean centring pre-processing to both calibrate and cross-validate the samples from the assigned regions. A *k*-fold cross-validation method ($k = 10$, Venetian blinds procedure) was selected considering the size of the sample set. Class predictions were based on classification rules "class predict strict" (probability > threshold value of 0.5) and "class predict most probable" (highest probability). Data compression and DA techniques were undertaken using Solo software.

3. Results and discussion

A rapid and robust analytical method was sought that could verify the categorical provenance of wine and provide a level of assurance within the supply chain and on to consumers. In the case of regional authentication of Australian wine, several studies have been undertaken (Bellomarino, Conlan, Parker, Barnett, & Adams, 2009; Liu, Cozzolino, Cynkar, Gishen, & Colby, 2006; Martin, Watling, & Lee, 2012; Riovanto, Cynkar, Berzaghi, & Cozzolino, 2011) using various techniques and with an emphasis on different wine varieties. Cabernet Sauvignon was considered for the present study, as it is the second most planted red grape variety in Australia after Shiraz, and not only features as a varietal in its own right in premium and regional wines, but also contributes importantly to red wine blends. As the birthplace of this varietal, predominant Cabernet Sauvignon blends from Bordeaux were included for a comparison.

3.1. Basic chemical analysis

Basic chemical parameters (pH, TA, alcohol content, wine colour measurements) determined for the Cabernet Sauvignon wines underwent one-way ANOVA according to region. Differences in pH, TA, TP, and TAn were statistically significant ($p < 0.05$) among the regions, whereas alcohol content and hue were not (Table S1 of the Supplementary material). Bordeaux Cabernet Sauvignon wines presented the highest pH and the highest total anthocyanin content compared to wines from the Australian regions. This could be explained by climatic and winemaking differences, where cooler climates tend to produce higher anthocyanin concentrations (Cozzolino, Cynkar, Dambergs, Gishen, & Smith, 2010), and acid adjustments to lower the pH being routine in Australia.

3.2. Multi-element analysis for regional classification

A direct dilution method was selected for the sample preparation over alternatives such as microwave digestion and filtration treatments considering the ease of use, accuracy, and precision of the approach (Godshaw, Hopfer, Nelson, & Ebeler, 2017). ANOVA revealed 25 elements, Li, Be, Na, Mg, Al, K, Ca, Sc, Ti, V, Mn, Co, Ni, Ga, As, Se, Rb, Sr, Mo, Cs, Ba, La, Ce, W, and Pb, differed significantly among the different regions with the means separated using Tukey's honestly significant difference test (Table S2 of the Supplementary material).

According to Table S2, elements such as K, Sc, Ti, V, As, La, W, and

Pb showed the highest least squares (LS) means for Bordeaux samples. Likewise, for Coonawarra region, Li, Na, Ca, and Se were found in the highest concentrations. In comparison, Ni for Margaret River and Mg for Bordeaux showed the lowest concentrations. Moreover, Sr has shown significantly different results among all four regions (highest in Coonawarra and lowest in Bordeaux), which is indicative of its relationship with the soil of origin as shown in other studies (Bora et al, 2017). Most of the significantly different elements found in the present work have been identified as responsible for geographical discrimination in several studies, as reported previously (Fan et al., 2018). Indeed, research conducted by Martin, Watling, and Lee (2012) on regional discrimination of Australian wines of different varieties including Cabernet Sauvignon showed variations in the concentrations of several elements in different regions, and similar to the present study, Na concentrations were higher in Coonawarra compared to other Australian regions. In contrast, Li concentration was highest in the Coonawarra wines of the present study whereas Martin, Watling, and Lee (2012) found Margaret River had the highest Li. This potentially highlighted the influence of other factors of wine production that can influence element concentrations (Waterhouse, Sacks, & Jeffery, 2016) and indicated the challenge of ascertaining a specific element responsible for the discrimination of individual regions.

3.2.1. Multivariate analysis with DA

Exploring the classification of Cabernet Sauvignon wines according to region, PCA was performed with significantly different elements to visualise any patterns in the data (Fig. 1a). Explained variance from the first two components was 39% and some clustering of regions was evident based on the influence of particular elements. For instance, Sr, Se, Li, Ca, and Na were characteristic of Coonawarra and Pb, Ti, K, Mo, Al, As, and Ce tended to define wines from Bordeaux. Similarly, higher levels of Cs, Rb and Ba were more associated with Yarra Valley whereas Mg and Mn were indicative of Margaret River. Further analysis of element data by DA after PCA reduction showed that F1 and F2 explained 82.40% of the total variance (Fig. 1b). The DA score plot shows a relatively clear separation of Bordeaux samples from the Australian wines, but there is some overlap between the Australian regions. The confusion matrix for the cross-validation results illustrates above 90% correct classification for each region, with Bordeaux showing 100% (Table 1). Two Coonawarra wines were misclassified as Margaret River, and only one each from Yarra Valley (as Margaret River) and Margaret River (as Coonawarra) were misclassified. While the results were encouraging using a common DA approach, a machine learning approach was investigated to try to improve the classification model.

Table 1
Confusion matrix for cross-validated DA and XGBDA models from ICP-MS data for the different regions (abbreviations as defined in the Materials and methods section).

	DA						XGBDA					
Predicted / Actual	B	C	M	Y	Total	Correct %	B	C	M	Y	Total	Correct %
B	10	0	0	0	10	100.00	10	0	0	0	10	100.00
C	0	33	2	1	36	91.67	0	35	0	1	36	97.20
M	0	1	19	0	20	95.00	0	0	20	0	20	100.00
Y	0	0	1	19	20	95.00	1	0	0	19	20	95.00
Total	10	34	22	20	86	94.19	11	35	20	20	86	97.67

3.2.2. Multivariate analysis with XGBDA

XGB has been identified as one of the most effective classification algorithms in machine learning in comparison to other methods (e.g. linear regression, decision tree, etc.), with the highest values in predictive performance in terms of sensitivity, specificity, accuracy, and area under receiver operating characteristics curve in cross-validation studies (Babajide Mustapha & Saeed, 2016). XGB modelling consists of a set of classification and regression trees (CART), implemented using the XGB algorithm:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F$$

where K = number of trees; f = function; F = all possible CARTs; x_i = descriptors of the given training set; y_i = class label. Accordingly, when each CART allocates a certain score to each recursive partition cell (leaf), the sum of prediction scores from each CART allows a unified approach for classification, which is further optimised in the process (Babajide Mustapha & Saeed, 2016).

Significantly different elements that were quantified by ICP-MS were reanalysed with XGBDA. Fig. 2a indicates the class prediction probability with XGBDA classification, calculated by considering the probabilities of each sample belonging to each possible class. Class prediction probability plots show results on a scale of 0 to 1 (0% –100%), where the sample values closer to 1 are more probable of being a member of the selected class. Also, all samples of the classes in Fig. 2a

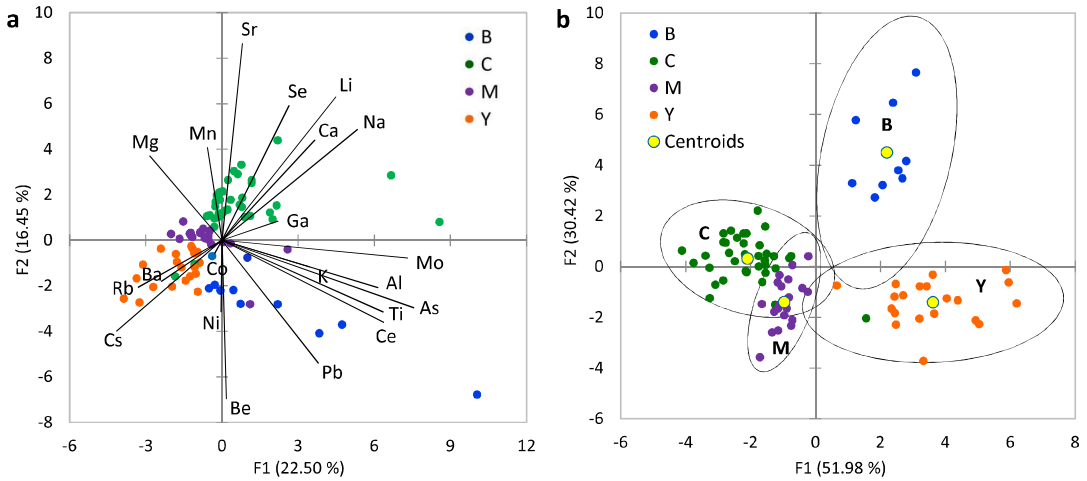


Fig. 1. (a) PCA biplot showing scores for the samples from the regions and loadings for the variables (significantly different elements) and (b) stepwise (forward) DA classification of regions based on PCA data; confidence ellipses correspond to a 95% confidence level. Region abbreviations: Coonawarra (C), Margaret River (M), Yarra Valley (Y), and Bordeaux (B).

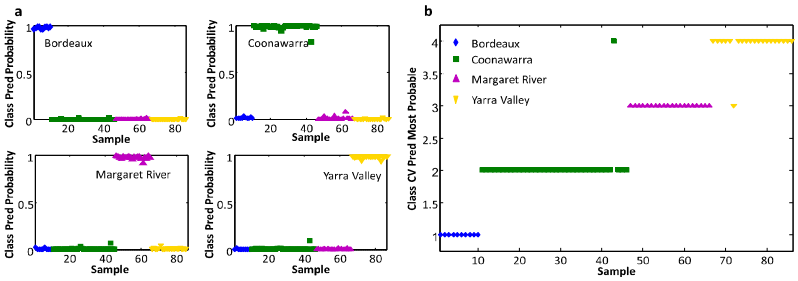


Fig. 2. XGBDA analysis of significant multi-element data from ICP-MS for wines of different regions showing (a) class prediction probability for each region and (b) class CV predicted most probable for all regions.

are above the strict threshold (0.5), meaning the model provides confidence that the sample belongs only to the assigned class according to the "class predict strict" classification rule. Class cross-validation (CV) predicted shown in Fig. 2b summarises the predicted classes for the samples using the Venetian blinds CV method, and assigns each sample to the class it most closely resembles ("class predict most probable" classification rule). Only two samples from two Australian regions were misclassified: Coonawarra (as Yarra Valley) and Yarra Valley (as Margaret River). The variable importance plot presented in Fig. S1 of the Supplementary material shows the measure of each variable's contribution to the model, with Cs, As, Mg, Al, Se and Sc being of most importance.

The evaluation of the confusion matrix (Table 1) shows somewhat higher prediction accuracy for XGBDA compared to the DA method described earlier, in the case of wines from Coonawarra (97.2% for XGBDA vs 91.7% for DA) and Margaret River (100% for XGBDA vs 95% for DA), and for overall total (97.7% vs 94.2%). Similarly, other evaluation parameters including sensitivity, specificity, precision and F1 scores (Tables S3a and b of the Supplementary material) gave higher values for XGBDA than DA for Coonawarra and Margaret River. Once again, Bordeaux was predicted with 100% accuracy using XGBDA, this time along with Margaret River, and only one sample was misclassified for each of Coonawarra and Yarra Valley.

In all cases, ICP-MS analysis of elements and modelling with DA and XGBDA showed better classification accuracy compared to the intraregional cross-validation results of 80% correct classification reported by Coetzee, van Jaarsveld, and Vanhaecke (2014). The present results approached the 100% correct classification level reported when considering Coonawarra, Yarra Valley and Great Southern (which includes Margaret River) among a broader study of Australian red wines (Martin, Watling, & Lee, 2012).

Having obtained very good modelling with the chosen chemometric approaches and ICP-MS data, attention was turned to fluorescence spectroscopy in an attempt to obtain the most robust and rapid classification technique (i.e., the combination of analytical and multivariate statistical methods).

3.3. A-TEEM analysis for regional classification

The Aqualog employed in this study uses right-angle optics to collect absorbances and correct for IFE, which would otherwise result in spectral distortion and a decrease of fluorescence intensities. Dilution of the sample can further help overcome IFE and helps standardise the sample matrix (Gilmore, 2014). Another factor to consider, especially in the case of phenolics, is the inhibition of oxidation. A study evaluating the effect of red wine oxidation by Gilmore et al. (2017) demonstrated the influence of oxidation on fluorescence intensity, revealing increased intensity across the absorbance spectrum and in the EEM. Avoidance of sample oxidation can be achieved by having a simple sample preparation procedure and rapid acquisition of spectral data, as was the case with the procedure employed (i.e., several minutes to prepare and analyse a sample).

3.3.1. EEM fingerprint

The EEM consists of signals obtained from fluorophores present in wine, with differences in intensities among samples due to the type and concentration of fluorophores. EEM surfaces of wine samples representing each region are presented as contour maps in Fig. 3, with each exhibiting a clear distinction in excitation/emission (EX/EM), especially at around 275/320 nm. The EEMs are unique for each individual wine and reflect a wine's molecular fingerprint, and features of the EEM data can be applied in classification of wine according to region or variety (Gilmore et al., 2017).

Airado-Rodríguez, Durán-Merás, Galeano-Díaz, and Wold (2011) summarised relevant fluorescent compounds naturally present in wine and their fluorescent properties (excitation wavelength (λ_{ex}) and emission wavelength (λ_{em})): for example, phenolic acids and aldehydes, λ_{ex} 260–320 and λ_{em} 320–440; flavonols, λ_{ex} 260–270, λ_{em} 370–420; monomeric and polymeric flavan-3-ols, λ_{ex} 280–290, λ_{em} 310–360. Variations of such compounds in wine will drive the differences in molecular fingerprint, making it possible to use these chemical markers (without necessarily having to identify individual analytes) for authentication of wine.

However, qualitative and quantitative visual interpretations of these differences are limited due to the overlapping of various signals from fluorophores (Gilmore et al., 2017) and a robust statistical analysis of these data is required for modelling of geographical origin. To extract relevant information from data, PARAFAC algorithms are generally applied for EEM as mentioned in the Introduction, although for the current data set consisting of unbalanced groups, component clustering did not reveal the correct classification due to overfitting of all the classes with the same loading (data not shown). Similarly, to the ICP-MS data, machine learning algorithms were explored with a view to improving the prediction model, this time assessing SVMDA along with XGBDA.

3.3.2. Multivariate analysis of EEM data with SVMDA

As a supervised machine learning algorithm, SVMDA is used to build non-linear models for classification. The unfolded EEM data were pre-processed with different options (see Section 2.5): SVMDA with PCA compression using 10 components and auto-scale pre-processing yielded the highest classification probability for each assigned class, using 95% confidence intervals (data not shown). Prediction accuracy for SVMDA with random split cross-validation ranged from a high of 90.0% for Bordeaux to a low of 72.5% for Yarra Valley, with a total correct classification of 84.7% (Table 2) being on par with other work.

However, the classification results were lower than the study of da Costa, Castro, and Barbosa (2016), who classified the geographical origin of Cabernet Sauvignon wines from Brazil and Chile with SVM and obtained 89% accuracy using antioxidant activity and phenolics

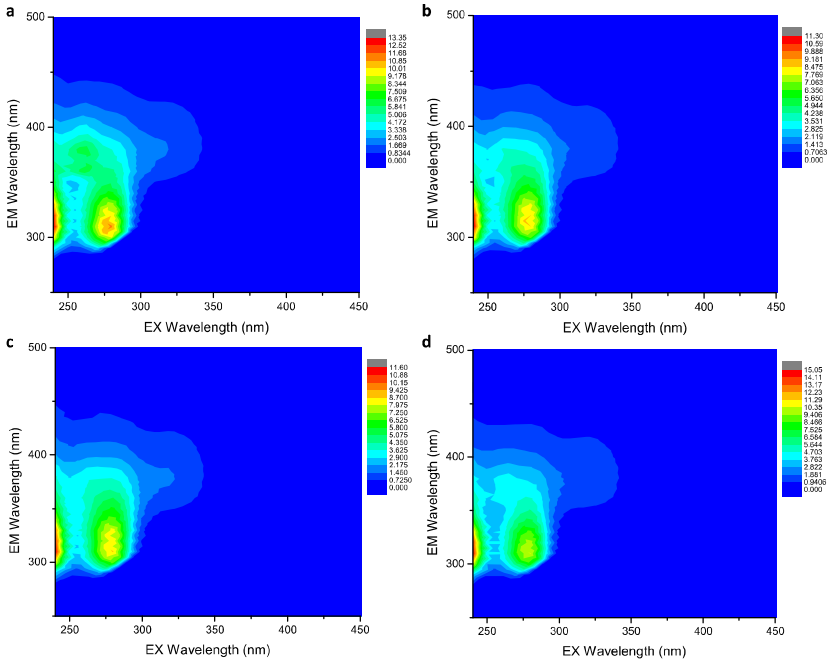


Fig. 3. EEM contour maps of wine samples from (a) Bordeaux, (b) Coonawarra, (c) Margaret River, and (d) Yarra Valley, after correction for the IFE, Rayleigh masking and normalisation according to water Raman scattering.

Table 2
Confusion matrix for SVMDA and XGBDA cross-validation results from EEM data for the different regions (Abbreviations as defined in the Materials and methods section).

	SVMDA						XGBDA					
Predicted / Actual	B	C	M	Y	Total[a]	Correct %	B	C	M	Y	Total[a]	Correct %
B	18	0	2	0	20	90.00	20	0	0	0	20	100.00
C	0	64	2	6	72	88.88	0	72	0	0	72	100.00
M	0	0	33	5	38	86.84	0	0	38	0	38	100.00
Y	0	7	4	29	40	72.50	0	0	0	40	40	100.00
Total	18	71	41	40	170	84.70	20	72	38	40	170	100.00

[a] Samples were diluted in duplicate and measured. One of the 86 wines was lost prior to dilution; hence, the total number of measurements was 170.

measures (with a balanced data set). Therefore, the present results were taken to indicate that SVMDA did not perform well in identifying the dimensions for correct classification with an unbalanced data set. The confusion matrix in Table 2 illustrates that all classes from the Australian regions overlapped with each other to some extent. However, none were misclassified as Bordeaux, thereby providing good differentiation of Cabernet Sauvignon wines from Australia and Bordeaux (recalling that the latter are blends with Cabernet Sauvignon as the dominant component). Furthermore, only two samples of Bordeaux were misclassified as Margaret River. Class prediction probability for SVMDA and class CV predicted further highlight the classifications (Fig. S2 of the Supplementary material), which were not as good as using multi-element data with XGBDA according to the classification rules.

3.3.3. Multivariate analysis of EEM data with XGBDA

The unfolded EEM data were analysed with the XGBDA algorithm that was described earlier. Quite remarkably, the score probabilities of XGBDA cross-validation indicated an unprecedented 100% correct classification of all tested wine regions using either rule for classification (Table 2). Class prediction probability and class CV predicted plots for XGBDA of EEM data (Fig. 4) show a distinct resolution among the different Cabernet Sauvignon producing regions, without any overlap between classes.

In comparison to using EEM data with SVMDA, it was clear that XGBDA provided a more selective classification, as was the case with ICP-MS data in comparison to DA. This can mainly be ascribed to the built-in capability for effective tree pruning and parallel processing of XGBDA (Babajide Mustapha & Saeed, 2016). Ultimately, the results confirmed the hypothesis regarding the successful application of fluorescence spectroscopy as a technique to build robust authentication models for wine.

3.4. Comparison of ICP-MS and EEM

The molecular fingerprinting capability and sensitivity of EEM data in conjunction with multivariate statistical analysis using the XGBDA algorithm provided sufficient chemical/spectral information to facilitate the completely accurate classification of Cabernet Sauvignon wines according to the region of origin. The performance of the XGBDA models using ICP-MS and EEM data sets in relation to sensitivity, specificity, precision and F1 values are shown in Tables S3b and c of the Supplementary material, highlighting the effectiveness of EEM data in correctly classifying the wines.

The 100% classification accuracy of EEM with XGBDA compared very favourably with multi-element studies such as Jurado, Alcázar, Palacios-Morillo, and de Pablos (2012) or Martin, Watling, and Lee (2012) for geographical authentication. Similarly, apart from a few misclassifications of the Australian Cabernet Sauvignons in the present work when using different chemometric models or multi-element data, results were readily comparable to the accuracy offered by other approaches. The misclassifications might be due to greater similarity among the terroir conditions and winemaking practices of the Australian regions, as reported previously in several studies (Liu, Cozzolino, Cynkar, Gishen, & Colby, 2006; Riovanto, Cynkar, Berzaghi, & Cozzolino, 2011). Overall, the results emphasised that Bordeaux

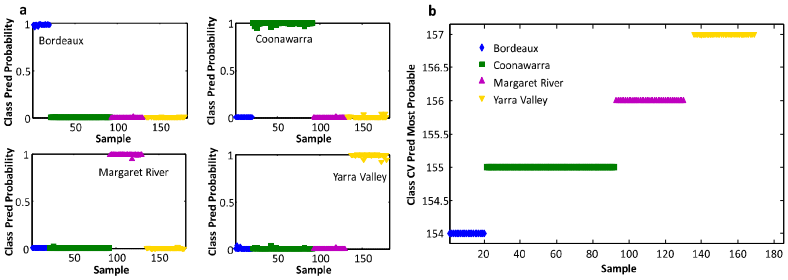


Fig. 4. XGBDA analysis of EEM data for wines of different regions showing (a) class prediction probability for each region (b) class CV predicted most probable for all regions.

samples were very readily distinguished from the Australian regions, as seen with the clear resolution upon classification using both ICP-MS and EEM results (Tables 1 and 2). Moreover, XGBDA modelling of EEM data was shown to be a very successful method for wine authentication according to geographical origin in comparison to modelling multi-element data.

4. Conclusions

Focusing on the most successful classification modelling, the XGBDA algorithm was used for classification of wine according to geographical origin. More specifically, the analysis of EEM data with XGBDA was identified as being very effective in the classification of these types of multiclass unbalanced groups. Notably, XGBDA revealed a higher capacity of prediction over the other classification models used with multi-element data, yielding 97.7% accuracy for significant elements determined by ICP-MS, and 100% using fluorescence data.

Moreover, the A-TEEM method was determined to be highly effective for the discrimination of commercial Cabernet Sauvignon wines from 2015 according to their geographical origins, with simpler sample preparation and less time-consuming analysis in comparison to ICP-MS or other elaborate authentication methods. Given the relative simplicity of A-TEEM, this approach could foreseeably be developed more broadly (e.g., for different varieties, vintages, and regions) and applied in the supply chain as a robust method for authentication or detection of adulterated wines. Ultimately, based on knowledge of the underlying chemical compounds leading to the molecular fingerprints, this methodology could also be beneficial for the optimisation of regional expression to improve the value of wines of provenance, and could be extended to other beverages or foods.

CRediT authorship contribution statement

Ranaweera K.R. Ranaweera: Conceptualisation, Formal analysis, Investigation, Methodology, Visualisation, Writing - original draft, Writing - review & editing. **Adam M. Gilmore:** Methodology, Formal analysis, Writing - review & editing. **Dimitra L. Capone:** Conceptualisation, Supervision, Writing - review & editing. **Susan E.P. Bastian:** Conceptualisation, Funding acquisition, Supervision, Writing - review & editing. **David W. Jeffery:** Conceptualisation, Funding acquisition, Project administration, Resources, Supervision, Writing - review & editing.

Declaration of Competing Interest

R.K.R.R., D.L.C., S.E.P.B. and D.W.J. declare no competing financial interest. A.M.G., as an employee of HORIBA Instruments Inc., has no competing financial interest to declare. HORIBA Instruments Inc. did not influence the interpretation of results or decision to publish the manuscript.

Acknowledgements

We are immensely grateful to Professor Andy Barker and his team at the University of New South Wales for assistance with Aqualog analysis. Dr Anthony Quinn, Quark Photonics Pty Ltd is thanked for facilitating access to the Aqualog instrument. We acknowledge the time and support provided for ICP-MS analysis by Dr Sarah Gilbert at Adelaide Microscopy, The University of Adelaide (a facility that is funded by the University, and State and Federal Governments), and are thankful to our UA colleagues for laboratory assistance. R.K.R.R. is appreciative of the financial support provided by an Australian Government Scholarship and Wine Australia supplementary scholarship (WA Ph1909). The Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au; project number IC170100008) is funded by the Australian Government with additional support from Wine Australia, Waite Research Institute, and industry partners. The University of Adelaide is a member of the Wine Innovation Cluster.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2020.127592.

References

Airado-Rodríguez, D., 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*(2), 257–264.

Al Riza, D. F., Widodo, S., Purwanto, Y. A., & Kondo, N. (2019). Authentication of the geographical origin of patchouli oil using front-face fluorescence spectroscopy and chemometric analysis. *Flavour and Fragrance Journal, 34*(1), 15–20.

Amargianitaki, M., & Spyros, A. (2017). NMR-based metabolomics in wine quality control and authentication. *Chemical and Biological Technologies in Agriculture, 4*, 9.

Azcarate, S. M., de Araújo Gomes, A., Alcaraz, M. R., Ugulino de Araújo, M. C., 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.

Babajide Mustapha, I., & Saeed, F. (2016). Bioactive molecule prediction using extreme gradient boosting. *Molecules, 21*(8), 983.

Bellomarino, S. A., Conlan, X. A., Parker, R. M., Barnett, N. W., & Adams, M. J. (2009). Geographical classification of some Australian wines by discriminant analysis using HPLC with UV and chemiluminescence detection. *Talanta, 80*(2), 833–838.

Bora, F.-D., Donici, A., Teodor, R., Bunea, A., Popescu, D., & Bunea, C.-I. (2017). Elemental profile and $^{207}Pb/^{206}Pb$, $^{208}Pb/^{206}Pb$, $^{204}Pb/^{206}Pb$, $^{87}Sr/^{86}Sr$ isotope ratio as fingerprints for geographical traceability of Romanian wines. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 46*, 223–239.

Coelho, C., Aron, A., Roullier-Gall, C., Gonsior, M., Schmitt-Kopplin, P., & Gougeon, R. D. (2015). Fluorescence fingerprinting of bottled white wines can reveal memories related to sulfur dioxide treatments of the must. *Analytical Chemistry, 87*(16), 8132–8137.

Coetzee, P. P., van Jaarsveld, F. P., & Vanhaecke, F. (2014). Intraregional classification of wine via ICP-MS elemental fingerprinting. *Food Chemistry, 164*, 485–492.

Cozzolino, D., Cynkar, W. U., Dambergs, R. G., Gishen, M., & Smith, P. (2010). Grape (*Vitis vinifera*) compositional data spanning ten successive vintages in the context of abiotic growing parameters. *Agriculture, Ecosystems & Environment, 139*(4), 565–570.

Cozzolino, D., Cynkar, W., Shah, N., & Smith, P. A. (2011). Can spectroscopy geographically classify Sauvignon Blanc wines from Australia and New Zealand? *Food*

Chemistry, 126(2), 673–678.
da Costa, N. L., Castro, I. A., & Barbosa, R. (2016). Classification of Cabernet Sauvignon from two different countries in South America by chemical compounds and support vector machines. *Applied Artificial Intelligence, 30*(7), 679–689.
Danezis, G. P., Tsagkaris, A. S., Camin, F., Brusic, V., & Georgiou, C. A. (2016). Food authentication: Techniques, trends & emerging approaches. *Trends in Analytical Chemistry, 85*, 123–132.
Dufour, É., Letort, A., Laguet, A., Lebecque, A., & Serra, J. N. (2006). Investigation of variety, typicality and vintage of French and German wines using front-face fluorescence spectroscopy. *Analytica Chimica Acta, 563*(1), 292–299.
Fan, S., Zhong, Q., Gao, H., Wang, D., Li, G., & Huang, Z. (2018). Elemental profile and oxygen isotope ratio ($\delta^{18}O$) for verifying the geographical origin of Chinese wines. *Journal of Food and Drug Analysis, 26*(3), 1033–1044.
Geana, E. I., Marinescu, A., Iordache, A. M., Sandru, C., Ionete, R. E., & Bala, C. (2014). Differentiation of Romanian wines on geographical origin and wine variety by elemental composition and phenolic components. *Food Analytical Methods, 7*(10), 2064–2074.
Gilmore, A. (2014). How to collect national institute of standards and technology (NIST) traceable fluorescence excitation and emission spectra. In Y. Engelborghs & A. J. W. G. Visser (Eds.), Fluorescence Spectroscopy and Microscopy: Methods and Protocols, (pp. 3-27). Totowa, NJ: Humana Press.
Gilmore, A., Akaji, S., & Csatorday, K. (2017). Spectroscopic analysis of red wines with A-TEEM molecular fingerprinting. *Readout, E49*, 41–48.
Godshaw, J., Hopfer, H., Nelson, J., & Ebeler, S. E. (2017). Comparison of dilution, filtration, and microwave digestion sample pretreatments in elemental profiling of wine by ICP-MS. *Molecules, 22*(10), 1609.
Grindlay, G., Mora, J., de Loos-Vollebregt, M. T. C., & Vanhaecke, F. (2014). Evaluation of the multi-element capabilities of collision/reaction cell inductively coupled plasma–mass spectrometry in wine analysis. *Talanta, 128*, 379–385.
Herbert, P., Barros, P., & Alves, A. (2000). Detection of port wine imitations by discriminant analysis using free amino acids profiles. *American Journal of Enology and Viticulture, 51*(3), 262–268.
Jackson, R. S. (2009). Nature and origins of wine quality. In R. S. Jackson (Ed.). *Wine Tasting* (pp. 387–426). (Second Edition). San Diego: Academic Press.
Jaitz, L., Siegl, K., Eder, R., Rak, G., Abranko, L., Koellensperger, G., & Hann, S. (2010). LC–MS/MS analysis of phenols for classification of red wine according to geographic origin, grape variety and vintage. *Food Chemistry, 122*(1), 366–372.
Jurado, J. M., Alcázar, Á., Palacios-Morillo, A., & de Pablos, F. (2012). Classification of Spanish DO white wines according to their elemental profile by means of support vector machines. *Food Chemistry, 135*(3), 898–903.
Kokkinofta, R., Fotakis, C., Zervou, M., Zoumpoulakis, P., Savvidou, C., Poulli, K., ... Kefalas, P. (2017). Isotopic and elemental authenticity markers: A case study on Cypriot wines. *Food Analytical Methods, 10*(12), 3902–3913.
Liu, L., Cozzolino, D., Cynkar, W. U., Gishen, M., & Colby, C. B. (2006). Geographic classification of Spanish and Australian Tempranillo red wines by visible and near-infrared spectroscopy combined with multivariate analysis. *Journal of Agricultural and Food Chemistry, 54*(18), 6754–6759.
Lohumi, S., Lee, S., Lee, H., & Cho, B.-K. (2015). A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science & Technology, 46*(1), 85–98.
Martin, A. F., Watling, R. J., & Lee, G. S. (2012). The multi-element determination and regional discrimination of Australian wines. *Food Chemistry, 133*(3), 1081–1089.
McLeod, R. (2017). *Counting the cost: Lost Australian food and wine export sales due to fraud.* Food Innovation Australia Ltd.
Mercurio, M. D., Dambergs, R. G., Herderich, M. J., & Smith, P. A. (2007). High throughput analysis of red wine and grape phenolics - adaptation and validation of methyl cellulose precipitable tannin assay and modified Somers color assay to a rapid 96 well plate format. *Journal of Agricultural and Food Chemistry, 55*(12), 4651–4657.
Obreque-Slier, E., Peña-Neira, Á., López-Solís, R., Zamora-Marín, F., Ricardo-da Silva, J. M., & Laureano, O. (2010). Comparative study of the phenolic composition of seeds and skins from Carménère and Cabernet Sauvignon grape varieties (*Vitis vinifera L.*) during ripening. *Journal of Agricultural and Food Chemistry, 58*(6), 3591–3599.
Pinu, F. R. (2018). Grape and wine metabolomics to develop new insights using untargeted and targeted approaches. *Fermentation, 4*(4), 92.
Quatela, A., Gilmore, A. M., Gall, K. E. S., Sandros, M., Csatorday, K., Siemiarczuk, A., ... Camenen, L. (2018). A-TEEM, a new molecular fingerprinting technique: Simultaneous absorbance-transmission and fluorescence excitation-emission matrix method. *Methods and Applications in Fluorescence, 6*(2), Article 027002.
Ríos-Reina, R., Elcoroaristizabal, S., Ocaña-González, J. A., García-González, D. L., Amigo, J. M., & Callejón, R. M. (2017). Characterization and authentication of Spanish PDO wine vinegars using multidimensional fluorescence and chemometrics. *Food Chemistry, 230*, 108–116.
Riovanto, R., Cynkar, W. U., Berzaghi, P., & Cozzolino, D. (2011). Discrimination between Shiraz wines from different Australian regions: The role of spectroscopy and chemometrics. *Journal of Agricultural and Food Chemistry, 59*(18), 10356–10360.
Suciu, R.-C., Zarbo, L., Guyon, F., & Magdas, D. A. (2019). Application of fluorescence spectroscopy using classical right angle technique in white wines classification. Scientific Reports, 9(1), 18250-18250.
Versari, A., Laurie, V. F., Ricci, A., Laghi, L., & Parpinello, G. P. (2014). Progress in authentication, typification and traceability of grapes and wines by chemometric approaches. *Food Research International, 60*, 2–18.
Villano, C., Lisanti, M. T., Gambuti, A., Vecchio, R., Moio, L., Frusciante, L., ... Carputo, D. (2017). Wine varietal authentication based on phenolics, volatiles and DNA markers: State of the art, perspectives and drawbacks. *Food Control, 80*, 1–10.
Waterhouse, A. L., Sacks, G. L., & Jeffery, D. W. (2016). *Understanding Wine Chemistry.* UK: John Wiley & Sons Ltd.
Yin, C., Li, H., Ding, C., & Wang, H. (2009). Preliminary investigation on variety, brewery and vintage of wines using three-dimensional fluorescence spectroscopy. *Food Science and Technology Research, 15*(1), 27–38.

SUPPLEMENTARY MATERIAL FOR

Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling

Ranaweera K. R. Ranaweera,[a] Adam M. Gilmore,[b] Dimitra L. Capone,[a,c] Susan E. P. Bastian,[a,c] David W. Jeffery*[,a,c]

[a] Department of Wine and Food Science, and Waite Research Institute, The University of Adelaide (UA), PMB 1, Glen Osmond, South Australia 5064, Australia

[b] HORIBA Instruments Inc. 20 Knightsbridge Rd., Piscataway, NJ 08854

[c] Australian Research Council Training Centre for Innovative Wine Production, UA, PMB 1, Glen Osmond, South Australia 5064, Australia

* Corresponding author.

E-mail address: david.jeffery@adelaide.edu.au (D.W. Jeffery).

Table of Contents

Page

Table S1. Least squares means and one-way ANOVA results of chemical measures within regions. S-2

Table S2. Least squares means and one-way ANOVA results of selected elements for the different regions. S-2

Table S3. Confusion matrices of different cross-validated models for the wine regions showing (a) DA of ICP-MS data, (b) XGBDA of ICP-MS data, and (c) XGBDA of EEM data. S-3

Figure S1. XGBDA variable score plot for ICP-MS multi-element data giving a measure of a variable's importance for building the model. S-3

Figure S2. SVMDA analysis of EEM data for wines of different regions showing (a) class prediction probability for each region and (b) class CV predicted for all regions. S-4

Table S1. Least squares means and one-way ANOVA results of chemical measures for the different regions: Coonawarra (C), Margaret River (M), Yarra Valley (Y), and Bordeaux (B).[a]

	C	Y	M	B	Pr > F(Model)
pH	3.56 c	3.59 bc	3.64 b	3.74 a	< 0.0001
TA	6.0 a	5.8 b	5.5 b	5.0 c	< 0.0001
Alcohol (% v/v)	14.2	14.0	14.0	13.8	0.19
Total average phenolics (au)	69 a	66 ab	56 c	62 b	< 0.0001
Total average anthocyanins (mg L^{-1})	126 b	120 b	130 b	169 a	0.007
Average hue	0.87	0.87	0.87	0.85	0.83

[a] Different letters within a row indicate statistically significant differences among the means according to Tukey HSD post hoc test (α = 0.05).

Table S2. Least squares means and one-way ANOVA results of selected elements for the different regions: Coonawarra (C), Margaret River (M), Yarra Valley (Y), and Bordeaux (B).[a]

Element	Units	C	Y	M	B	Pr > F(Model)
Li	µg L^{-1}	1.4 a	0.7 b	1.0 b	1.0 b	< 0.0001
Be	µg L^{-1}	0.03 b	0.06 a	0.03 b	0.05 a	< 0.0001
Na	mg L^{-1}	3.9 a	1.3 c	2.6 b	1.8 bc	< 0.0001
Mg	mg L^{-1}	20.2 b	21.5 a	20.6 ab	16.1 c	< 0.0001
Al	µg L^{-1}	38.9 ab	29.2 b	49.3 a	53.0 a	0.002
K	mg L^{-1}	105.1 b	106.4 b	101.6 b	124.8 a	0.002
Ca	mg L^{-1}	10.5 a	9.2 b	9.2 b	8.8 b	0.002
Sc	µg L^{-1}	0.01 b	0.01 b	0.01 b	0.03 a	0.001
Ti	µg L^{-1}	7.4 b	6.8 b	6.8 b	9.9 a	0.002
V	µg L^{-1}	0.9 b	1.2 b	0.64 b	17.5 a	< 0.0001
Mn	µg L^{-1}	256.7 a	235.4 ab	184.4 bc	155.4 c	0.000
Co	µg L^{-1}	0.56 b	0.72 a	0.47 b	0.46 b	0.010
Ni	µg L^{-1}	2.7 b	4.1 a	1.72 c	4.1 a	< 0.0001
Ga	µg L^{-1}	0.03 b	0.03b	0.05 a	0.03 ab	0.000
As	µg L^{-1}	0.5 b	0.22 b	0.21 b	1.1 a	< 0.0001
Se	µgL^{-1}	0.5a	0.24 b	0.21 b	0.31 b	< 0.0001
Rb	µg L^{-1}	329.2 b	409.9 a	472.4 a	304.9 b	< 0.0001
Sr	µg L^{-1}	222.0 a	112.0 c	169.5 b	42.2 d	< 0.0001
Mo	µg L^{-1}	0.70 ab	0.25 c	0.40 bc	0.99 a	0.005
Cs	µg L^{-1}	0.4 b	2.3 a	1.8 a	0.78 b	< 0.0001
Ba	µg L^{-1}	30.3 b	66.5 a	19.4 c	20.0 bc	< 0.0001
La	µg L^{-1}	0.05 b	0.02 b	0.06 b	0.16 a	0.021
Ce	µg L^{-1}	0.07 b	0.03 b	0.11 ab	0.23 a	0.035
W	µg L^{-1}	0.16 b	0.23 b	0.10 b	0.63 a	0.009
Pb	µg L^{-1}	0.78 b	0.71 b	0.66 b	2.4 a	< 0.0001

[a] Different letters within a row indicate statistically significant differences among the means according to Tukey HSD post hoc test (α = 0.05).

Table S3. Confusion matrices showing the performance parameters of different cross-validated models for the wine regions showing (a) DA of ICP-MS data, (b) XGBDA of ICP-MS data, and (c) XGBDA of EEM data.

a

Region	Sensitivity	Specificity	Precision	F1 Score
Bordeaux	1.0	1.0	1.0	1.0
Coonawarra	0.92	0.94	0.91	0.94
Margaret River	0.95	0.98	0.95	0.90
Yarra Valley	0.95	0.98	0.95	0.95

b

Region	Sensitivity	Specificity	Precision	F1 Score
Bordeaux	1.0	1.0	1.0	1.0
Coonawarra	0.97	1.0	1.0	0.98
Margaret River	1.0	0.98	0.95	0.97
Yarra Valley	0.95	0.98	0.95	0.95

c

Region	Sensitivity	Specificity	Precision	F1 Score
Bordeaux	1.0	1.0	1.0	1.0
Coonawarra	1.0	1.0	1.0	1.0
Margaret River	1.0	1.0	1.0	1.0
Yarra Valley	1.0	1.0	1.0	1.0

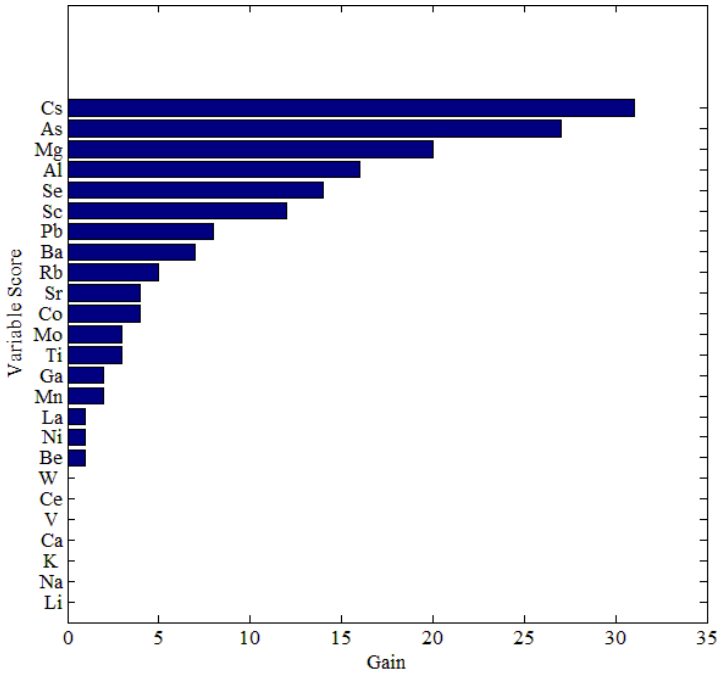


Fig. S3. XGBDA variable score plot for ICP-MS multi-element data giving a measure of a variable's importance for building the model.

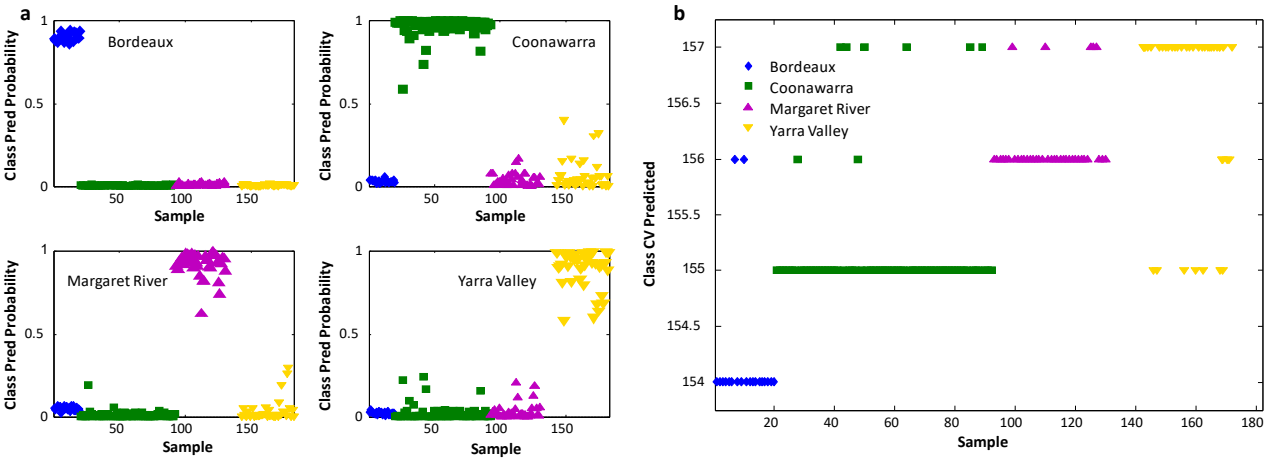


Fig. S4. SVMDA analysis of EEM data for wines of different regions showing (a) class prediction probability for each region and (b) class CV predicted for all regions.

Chapter 4

Publication

Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. Food Chemistry, 361, 130149.

Statement of Authorship

Title of Paper	Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. Food Chemistry, 361, 130149. https://doi.org/https://doi.org/10.1016/j.foodchem.2021.130149

Principal Author

Name of Principal Author (Candidate)	Ranaweera K. R. Ranaweera		
Contribution to the Paper	Conceived of and designed the experiments. Performed analysis of basic chemical parameters and A-TEEM analysis for the wine samples. Carried out data collection, statistical analysis and interpretation of data for the manuscript. Drafted the manuscript, and reviewed and edited the final draft for publication.		
Overall percentage (%)	65%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/07/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Adam M. Gilmore		
Contribution to the Paper	Contributed to building regression models for phenolic quantification of wine samples and assisted with statistical techniques in formal analysis of the study data. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	2021/08/03

Name of Co-Author	Dimitra L. Capone		
Contribution to the Paper	Contributed to the conception and design of experiments. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	20/07/2021

Name of Co-Author	Susan E.P. Bastian		
Contribution to the Paper	Contributed to the conception and design of experiments. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	19.7.2021

Name of Co-Author	David W. Jeffery		
Contribution to the Paper	Original conceptualisation of the study. Acquired funding and provided resources for the project leading to this publication. Supervised the project and undertook project administration throughout the study. Reviewed, edited and revised the manuscript for publication. Acted as the corresponding author at all stages.		
Signature		Date	19/07/2021

Food Chemistry 361 (2021) 130149

Contents lists available at ScienceDirect

Food Chemistry

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

Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine

Ranaweera K.R. Ranaweera [a], Adam M. Gilmore [b], Dimitra L. Capone [a,c], Susan E.P. Bastian [a,c], David W. Jeffery [a,c,*]

[a] Department of Wine Science and Waite Research Institute, The University of Adelaide (UA), PMB 1, Glen Osmond, South Australia 5064, Australia
[b] HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, United States
[c] Australian Research Council Training Centre for Innovative Wine Production, UA, PMB 1, Glen Osmond, South Australia 5064, Australia

ARTICLE INFO

Keywords:
Extreme gradient boosting
Polyphenols
Multi-block data
Authenticity
Chemometrics
Vitis Vinifera

ABSTRACT

Fluorescence spectroscopy is rapid, straightforward, selective, and sensitive, and can provide the molecular fingerprint of a sample based on the presence of various fluorophores. In conjunction with chemometrics, fluorescence techniques have been applied to the analysis and classification of an array of products of agricultural origin. Recognising that fluorescence spectroscopy offered a promising method for wine authentication, this study investigated the unique use of an absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) technique for classification of red wines with respect to variety and geographical origin. Multi-block data analysis of A-TEEM data with extreme gradient boosting discriminant analysis yielded an unrivalled 100% and 99.7% correct class assignment for variety and region of origin, respectively. Prediction of phenolic compound concentrations with A-TEEM based on multivariate calibration models using HPLC reference data was also highly effective, and overall, the A-TEEM technique was shown to be a powerful tool for wine classification and analysis.

1. Introduction

The application of spectrofluorometry is of interest for food analysis because of the high sensitivity, selectivity, and rapidity of the technique. Based on the presence of fluorophores such as phenolic compounds, vitamins, and aromatic amino acids, spectrofluorometry has been combined with multivariate statistical analysis and used for qualitative investigation of food products (e.g., dairy, meat, egg, fruits, vegetables) to discriminate, classify, or identify adulteration, as well as for quantitative measurements (Karoui & Blecker, 2011). For a beverage like wine, which is considered a luxury good and therefore prone to fraudulent activity, the abundance of fluorophoric compounds means that fluorescence spectroscopy offers great potential as an analytical tool for authentication and analyte quantification.

Fluorescence spectroscopy has been applied in wine analysis to examine oenological phenomena related to oxidation or sulfur dioxide addition (Coelho et al., 2015; Gilmore, Akaji, & Csatorday, 2017), for quantification of polyphenols (Cabrera-Bañegil, Hurtado-Sánchez, Galeano-Díaz, & Durán-Merás, 2017), and for discrimination according to geographical origin, quality grade or grape variety (Dufour, Letort, Laguet, Lebecque, & Serra, 2006; Ranaweera, Gilmore, Capone, Bastian, & Jeffery, 2021; Saad, Bouveresse, Locquet, & Rutledge, 2016; Sádecká & Jakubíková, 2020; Suciu, Zarbo, Guyon, & Magdas, 2019). In the case of authentication, rapid spectroscopic methods are contrasted against more involved chemical measures of elemental composition, stable isotope ratios, amino acid profile, grape and wine volatile compounds, and polyphenols, arising from a range of advanced analytical techniques (Ranaweera, Souza Gonzaga, Capone, Bastian, & Jeffery, 2021). Fluorescence spectroscopy is appealing because of specific fluorophore compounds in wine providing a molecular fingerprint that is unique for each wine (Gilmore et al., 2017). This fingerprint consists of a three-dimensional excitation and emission matrix (EEM) formed by various fluorophores in the sample (Coelho et al., 2015).

Phenolic compounds are the dominant fluorescent molecules in wine

* Corresponding author at: Department of Wine Science and Waite Research Institute, The University of Adelaide, PMB 1, Glen Osmond, South Australia 5064, Australia.
E-mail address: david.jeffery@adelaide.edu.au (D.W. Jeffery).

https://doi.org/10.1016/j.foodchem.2021.130149
Received 16 February 2021; Received in revised form 21 April 2021; Accepted 15 May 2021
Available online 18 May 2021

(Airado-Rodríguez, Durán-Merás, Galeano-Díaz, & Wold, 2011) and are responsible for most of the biological activity of wine because of their antioxidant, anti-inflammatory and anti-microbial properties (Cabrera-Bañegil et al., 2017). Moreover, they are one of the major determinants of red wine quality in regard to colour, mouthfeel and taste. Phenolic compounds are classified as non-flavonoids, which includes hydroxycinnamic acids, hydroxybenzoic acids, hydrolysable tannins, and stilbenes, or as flavonoids, which comprises monomeric and polymeric flavan-3-ols, flavonols, and anthocyanins (Waterhouse, 2002). The phenolic composition of wine can be influenced by factors including grape variety, viticultural and oenological practices, and physical phenomena such as climate, light, and soil condition in the vineyard (Waterhouse, Sacks, & Jeffery, 2016). Therefore, grape and wine phenolics have been considered as a prominent chemical marker for wine authentication in the past using various analytical methods (Vergara, von Baer, Mardones, & Gutiérrez, 2011). Quantitative and qualitative analysis of phenolic compounds has mostly been undertaken with high-performance liquid chromatography (HPLC), which affords precise measurements (Ferrer-Gallego, Rodríguez-Pulido, Toci, & García-Estevez, 2020). However, chromatographic techniques can be very time consuming and involve relatively complicated sample preparation and analytical procedures; fluorescence spectroscopy offers an advantageous alternative (Cabrera-Bañegil et al., 2017).

The fluorescence technique can be categorised as conventional, total EEM luminescence spectroscopy, and synchronous spectrofluorometry (Dankowska, 2016). The conventional approach involves keeping the excitation wavelength (λ_{ex}) or emission wavelength (λ_{em}) constant while scanning the other, which is problematic for identifying certain analytes in a multicomponent sample (Samokhvalov, 2020). In contrast, the synchronous method simultaneously scans both the λ_{ex} and λ_{em} at a constant difference of $\Delta\lambda = \lambda_{em} - \lambda_{ex}$, which is more applicable in quantitative and qualitative analysis (Samokhvalov, 2020). In the total EEM method, a set of emission spectra is recorded within a set value of λ_{ex}, enabling full information on the fluorescent species present in the sample to be obtained, which is most advantageous for pattern recognition among samples that differ only slightly in their composition (Airado-Rodríguez et al., 2011). In addition, according to the geometry of sample illumination, the fluorescence technique can be identified as right-angle or front-face, which dictates the type of sample and data preparation that needs to be performed, e.g., dilution of sample for right-angle analysis and normalisation of spectra for the front-face technique (Dankowska, 2016). Dilution of a food sample can make it difficult to extrapolate the results to the original food matrix (Kulmyrzaev, Karoui, De Baerdemaeker, & Dufour, 2007), although dilution is frequently used in spectrophotometric studies of wine (Aleixandre-Tudo, Buica, Nieuwoudt, Aleixandre, & du Toit, 2017). Comparing different fluorescence techniques for brandy classification according to region of origin, Sádecká, Uríčková, Májek, and Jakubíková (2019) showed that diluted samples recorded with right-angle geometry provided more accurate classification than the front-face method.

Red wine spectral measures typically involve certain UV–vis wavelengths (e.g., 280, 420, 520 nm), so an absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) technique has been developed using an Aqualog instrument with fluorescence detection via right-angle optical geometry (Gilmore et al., 2017). This approach yields spectral information in the UV–vis range for all chromophores and fluorophores, thus giving EEM fingerprints in combination with absorbance data. Advantageously, fusion of such data has been shown by Tan et al. (2016) to considerably improve the accuracy of classification. Whatever the approach, the application of multivariate statistical methods is important in extracting qualitative and quantitative information because of the overlap and variable intensity of signals from the fluorescent compounds in wine (Gilmore et al., 2017).

In multivariate data analysis, unsupervised decomposition algorithms such as principal component analysis (PCA) are important in the visualisation of data. A generalisation of two-way PCA, parallel factor analysis (PARAFAC) is commonly applied to EEM data as it specifically explains the variability of all EEMs in a trilinear decomposition model (Cabrera-Bañegil, Hurtado-Sánchez, Galeano-Díaz, & Durán-Merás, 2017). Several other methods such as independent components analysis (Saad et al., 2016), factorial discriminant analysis (Dufour et al., 2006), and partial least squares (PLS) discriminant analysis (Azcarate et al., 2015) have been applied in wine authentication with over 80% correct classification. More recently, a machine learning technique known as extreme gradient boosting (XGB) discriminant analysis was applied with 100% effectiveness in a classification study of regional Cabernet Sauvignon wines using EEM data (Ranaweera et al., 2021). The advantage of XGB is that it is based on a decision tree model and implements gradient boosting in an optimised setting, where the algorithm combines multiple pruned trees of low accuracy and reduces the error in consequent iterations (Chen & Guestrin, 2016). Other than classification, multivariate techniques are applied in calibration/quantification, developing regression models such as PLS regression (PLSR) or principal component regression using analyte concentrations determined for a set of reference samples (Jiménez-Carvelo, González-Casado, Bagur-González, & Cuadros-Rodríguez, 2019). The developed models can be applied for the prediction of component concentration in unknown samples based on their fluorescence spectra.

With the need for a robust but simple authentication method for use within the production and supply chain, we tested the hypothesis that the A-TEEM technique in combination with chemometrics could be applied to the classification of red wines from different Australian Geographical Indications (GI) according to their regional and varietal origins. The effectiveness of the cross-validated modelling technique was evaluated according to the score probabilities in the confusion matrix. Additionally, the A-TEEM data was used for predicting the concentrations of 24 phenolic compounds in the set of wines according to PLSR modelling with HPLC reference data, utilising PCA to examine the relationship between the regions or varieties and the phenolic profiles.

2. Material and methods

2.1. Chemicals and solvents

HPLC gradient grade absolute ethanol, analytical grade 37% hydrochloric acid (HCl), and sodium hydroxide pellets (98%) were purchased from Chem-Supply (Port Adelaide, SA, Australia). High purity water was obtained through a Milli-Q purification system (Millipore, North Ryde, NSW, Australia).

2.2. Wine samples

A total of 221 samples of unreleased (2019 vintage) commercial Shiraz, Cabernet Sauvignon, and Merlot wines from ten different Australian regions (Barossa Valley, Clare Valley, Eden Valley, Langhorne Creek, McLaren Vale, Riverland, Wrattonbully, Murray Darling, Margaret River and Frankland River) were provided by wineries within the Accolade Wines portfolio (Table S1 and Fig. S1 of the Supplementary material).

2.3. Analytical procedures for determination of pH, TA and alcohol

Wine pH and titratable acidity (TA) measurements were obtained with a T50 autotitrator (Mettler Toledo, Melbourne, VIC, Australia), and ethanol content (percentage alcohol by volume, % v/v) was measured with an Alcolyzer Wine M/ME (Anton Paar, Graz, Austria).

2.4. Analytical procedure for A-TEEM

Wine samples were analysed in duplicate by the A-TEEM technique according to Ranaweera, Gilmore et al. (2021) with slight modifications.

Samples were diluted with 50% aqueous ethanol adjusted with HCl to pH 2, with a wine dilution factor of 1:150 determined by considering a linear relationship of fluorescent compounds versus absorbances according to the Beer-Lambert law (Gilmore, 2014). The solvent was vacuum filtered (0.45 μm PTFE membrane) and wine samples were centrifuged at 9300 × *g* for 10 min before dilution. Samples were analysed in a Hellma type 1FL (1 cm path length) Macro Fluorescence cuvette (Sigma-Aldrich, Castle Hill, NSW, Australia) within the excitation wavelength range of 240–700 nm with a 5 nm increment under medium gain and 0.2 s integration time. The emission wavelength range of 242–824 nm with a 4.66 nm increment as set by the instrument. An Aqualog spectrophotometer (version 4.2, Horiba Scientific, Quark Photonics, Adelaide, SA, Australia) was used to record the absorbance spectra (240–700 nm) and EEMs of the samples with data acquisition undertaken with Origin software (version 8.6, OriginLab Corporation, Massachusetts, USA). EEMs from the analysis of duplicate samples were separately pre-processed prior to statistical analysis by normalising according to the water Raman scattering units for the specified emission conditions and correcting for the influence of inner filter effects (IFE), solvent background, dark detector signals, and Rayleigh masking to eliminate spectral distortion (Gilmore et al., 2017).

2.5. Analytical procedure for phenolic compounds quantification

Phenolic concentrations were predicted using calibration models developed previously with HPLC reference values (Gilmore, Schober, Penichet, Zincker, & González, 2020). The following 24 phenolic compounds from different classes were included: anthocyanins (cyanidin-3-*O*-glucoside, delphinidin-3-*O*-glucoside, malvidin-3-*O*-glucoside, malvidin-3-*O*-acetylglucoside, malvidin-3-*O*-coumaroylglucoside, peonidin-3-*O*-glucoside, peonidin-3-*O*-acetylglucoside, peonidin-3-*O*-coumaroylglucoside, petunidin-3-*O*-glucoside), flavan-3-ols (catechin, epicatechin), flavonols (myricetin-3-*O*-galactoside, myricetin-3-*O*-glucoside, myricetin, quercetin-3-*O*-galactoside, quercetin 3-*O*-glucoside, quercetin, laricitin-3-*O*-glucoside, syringetin-3-*O*-galactoside, syringetin-3-*O*-glucoside, kaempferol), and hydroxycinnamates (caftaric acid, caffeic acid, coutaric acid). Phenolic compound concentrations (and totals per compound class) were correlated using PLSR with A-TEEM data (absorbance and EEM) obtained from Aqualog analysis.

Briefly, HPLC analyses were conducted according to Gilmore et al. (2020) with a 1290 Infinity Series HPLC equipped with diode array detector, thermostatted column compartment and autosampler (Agilent Technologies Santa Clara, USA). Separation occurred with a Zorbax SB C18 column (150 × 2.1 mm, 5 μm, Agilent Technologies, Santa Clara, USA) using mobile phases consisting of 87:10:3 v/v/v water:formic acid: acetonitrile (mobile phase A) and 40:10:50 v/v/v water:formic acid: acetonitrile (mobile phase B). A 20 μL injection volume was used except for malvidin-3-*O*-glucoside, which was 10 μL.

Anthocyanins: Mobile phase flow rate was 0.2 mL with the following gradient: 0–6 min, 4% B; 6–13 min, 30% B; 13–15 min, 40% B; 15–18 min, 45% B; 18–23 min, 30% B; 23–25 min, 4% B. Column temperature was 40 °C and detection occurred at 518 nm.

Other phenolics: Mobile phase flow rate was 0.5 mL/min with the following gradient: 0–15 min 6% B; 15–30 min, 32% B; 30–35 min, 42% B; 35–40 min, 50% B; 40–45 min, 80% B; 45–55 min, 65% B; 55–65 min, 6% B. Column temperature was 40 °C and detection occurred at 280 nm (flavan-3-ols), 320 nm (hydroxycinnamic acids), and 360 nm (flavonols).

2.6. Statistical analyses

Phenolic concentrations, pH, TA, and % alcohol by volume (v/v) were analysed by one-way analysis of variance (ANOVA) and Games-Howell (unequal group sizes) post-hoc test for pairwise comparisons ($\alpha = 0.05$), and PCA was carried out to explore the differences according to region and variety, using XLSTAT (version 2019.03.02, Addinsoft, Boston, USA). EEM data were unfolded into a two-way array using transform unfold multiway (mode 1) and analysed individually (EEM only) as well as with multi-block data modelling in combination with absorbance spectra, using Solo software (version 8.7.1, Eigenvector Research, Inc., Manson, WA, USA). Absorbance variables were baseline adjusted and the multi-block data was adjusted for dilution factor and integration time. For classification of samples according to region and variety, XGB discriminant analysis (XGBDA) was applied with PLS compression, using a maximum of 25 latent variables (LV), with mean centring pre-processing and decluttering with generalised least squares weighting (GLSW) at 0.2 to both calibrate and cross-validate (k = 10, Venetian blinds procedure). The effectiveness of the model was evaluated by the confusion matrix score probabilities, which are calculated from the values for true positive (TP), false positive (FP), true negative (TN) and false negative (FN) for each class, according to the "most probable" prediction rule of the given classification algorithm (Eigenvector, 2017). This enabled the calculation of the following:

Sensitivity (true positive rate): proportion of positive cases that were correctly identified = 100 × TP/(TP + FN);

Specificity (true negative rate): proportion of negative cases that were classified correctly = 100 × TN/(TN + FP);

Misclassification error: proportion of samples which were incorrectly classified = 100×(1-accuracy), where accuracy = (TP + TN)/(TP + TN + FP + FN);

Precision: proportion of positive cases giving a true positive result = 100 × TP/(TP + FP);

F1 Score: harmonic mean of precision and sensitivity = 2TP/(2TP + FP + FN).

Values of each parameter are within the range of 0 to 1 and can be presented as percentages, except for F1 score. PARAFAC analysis was used after imposing non-negativity constraints in all modes to explore the variability of the EEM data and identify the main fluorophores present in the sample. Calibration models were constructed with the PLSR algorithm using A-TEEM data and HPLC reference data to predict the phenolic compound concentrations. This involved a total of 131 commercial and experimental wines analysed in triplicate with 90:10 calibration/validation split for building the model. Optimisation of the model was based on the optimal prediction coefficient of determination as a function of the PLS latent variables. Model performance was assessed according to the ratio of root mean square error of cross-validation (RMSECV) (k = 10, Venetian blinds) and root mean square error of calibration (RMSEC) along with coefficient of determination (R^2 adjusted) and standard error of the regression slope (SE). Pearson's correlation (r) was used to measure the effectiveness of correlation. Predicted phenolic concentrations were examined with Solo software for variation based on region and variety, using robust PCA with GLSW.

3. Results and discussion

3.1. Basic chemical parameters of wines

The basic chemical parameters of varietal red wines from 10 different Australian GI were first explored. Average values for pH, TA and % v/v alcohol of the wines by variety and region are presented in Tables S2 and S3 of the Supplementary material are typical for Australian red wines (Godden, Wilkes, & Johnson, 2015). The data were subjected to one-way ANOVA and Games-Howell post-hoc test, showing that pH, TA and alcohol content differed significantly for region and variety. It is well-known that the various aspects of terroir such as grape variety, soil, climate, and viticultural practices can have a certain impact on the development and composition of grapes, which along with oenological inputs ultimately influence the characters of wine (Van Leeuwen & Seguin, 2006). SA regions tended to have higher pH, TA and alcohol values compared to the WA regions (Frankland River and Margaret River), but differences in basic chemistry did not discriminate the wines according to variety or region based on PCA (data not shown).

3.2. A-TEEM analysis

The use of EEM data was identified in a previous study as a rapid and robust analytical approach for verifying the provenance of commercial Cabernet Sauvignon from three GI of Australia and Bordeaux, France (Ranaweera, Gilmore et al., 2021). The current study introduced different varieties and regions, and the absorbance and fluorescence data from A-TEEM were combined into a multi-block data set to maximise robustness, as undertaken by Gilmoreet al. (2020).

3.2.1. Regional classification with XGBDA of A-TEEM data

The combined absorbance and EEM data were analysed according to wine region of origin with XGBDA. PLS compression was undertaken on pre-processed multi-block data to make the XGBDA model more stable and less prone to overfitting, thereby improving sensitivity. Fig. 1 indicates the class cross-validation (CV) prediction probability with XGBDA classification using the Venetian blinds CV method, which calculates the probability of each sample belonging to the class it most closely resembles. The classification performance of XGBDA models for each class was assessed from the confusion matrices obtained in cross-validation. The results are shown in Table 1, which outlines sensitivity, specificity, error, precision, and F1 score. Highlighting the power of this machine learning algorithm for classification, the score probabilities of XGBDA CV showed a remarkable result, with only one sample replicate (Fig. 1, circled) from McLaren Vale being misclassified as Riverland among > 200 samples analysed in duplicate.

According to Table 1, sensitivity (proportion of correctly classified positives) was 100% for all the classes except for McLaren Vale, which was 98%. Specificity (proportion of correctly classified negatives) of all the classes was also 100% except for Riverland, which was > 99%. Overall, the model showed very good prediction ability, with high precision and F1 score and a small error rate of 0.23% for Riverland and McLaren Vale classes because of the misclassification of one sample out of 442. Compared to using EEM alone (3 misclassifications out of 442, Table S4 of the Supplementary material), the combination of absorbance data with EEM slightly improved the score probabilities for regional classification. Compared to the data fusion method (synchronous fluorescence and UV–vis data) of Tan et al. (2016), which yielded 79.2% correct classification according to geographical origin of Chinese Cabernet Sauvignon wine, the results of A-TEEM technique in the present study showed an excellent outcome, achieving 99.7% correct classification overall. This was reminiscent of the accuracy achieved previously with EEM data and XGBDA modelling (Ranaweera et al., 2021) and compared very well to studies conducted to differentiate the geographical origin of Australian wine using other analytical methods. For example, Martin, Watling, and Lee (2012) applied inductively coupled plasma-mass spectrometry (ICP-MS) for differentiation of wine regions and correctly classified wines according to different states (Western Australia, Victoria, and South Australia), but had difficulty in separating wines within South Australian regions. Similarly, Riovanto, Cynkar, Berzaghi, and Cozzolino (2011) applied several spectroscopic techniques (UV–vis, near-infrared, and mid-infrared) with chemometrics and successfully discriminated Shiraz wines produced in Western Australia from South Australia, but yielded several misclassifications among South Australian regions. Such misclassification has been

Table 1
Confusion matrix results for XGBDA analysis of multi-block data (absorbance + EEM) for wines (in duplicate) according to GI.

Region	N	Sensitivity %	Specificity %	Error %	Precision %	F1 Score
Barossa Valley	56	100.00	100.00	0.00	100.00	1.00
Clare Valley	16	100.00	100.00	0.00	100.00	1.00
Eden Valley	24	100.00	100.00	0.00	100.00	1.00
Frankland River	14	100.00	100.00	0.00	100.00	1.00
Langhorne Creek	50	100.00	100.00	0.00	100.00	1.00
Margaret River	34	100.00	100.00	0.00	100.00	1.00
McLaren Vale	50	98.00	100.00	0.23	100.00	0.99
Riverland	168	100.00	99.64	0.23	99.41	0.99
Wrattonbully	20	100.00	100.00	0.00	100.00	1.00
Murray Darling	10	100.00	100.00	0.00	100.00	1.00

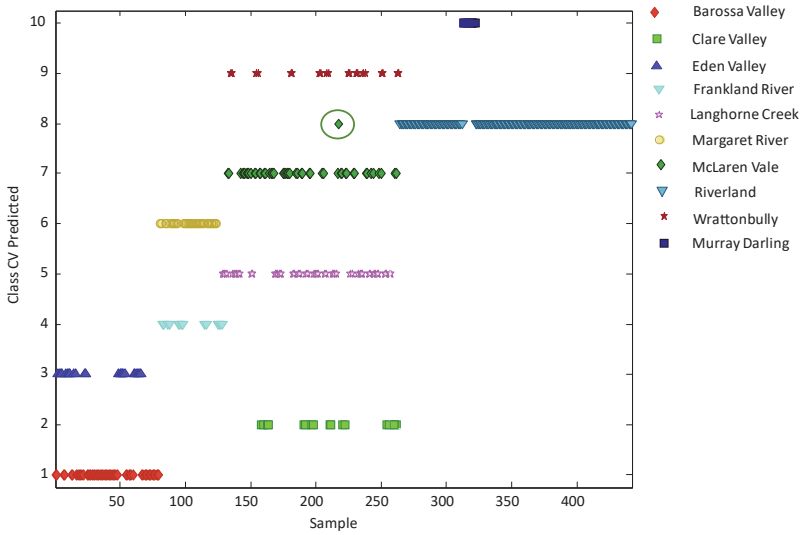


Fig. 1. Class CV predicted by region from XGBDA analysis of A-TEEM data for varietal red wines (n = 221, analysed in duplicate) from 10 different GIs of Australia. Only one sample (circled) was misclassified. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

explained on the basis of the influence of similar geology and climate across wine regions as a result of their proximity (Martin et al., 2012; Riovanto et al., 2011), although perhaps the techniques are simply not sensitive enough. In contrast to those studies, the present results showed that the A-TEEM technique combined with machine learning modelling could not only readily differentiate wine from GIs at a state level, (i.e., several thousand km apart) but also GIs within a state and even those that are adjacent (such as Barossa, Clare and Eden Valleys in South Australia) as seen from the results in Table 1 and Fig. 1.

3.2.2. Varietal classification with XGBDA of A-TEEM data

Varietal discrimination with A-TEEM was undertaken similarly to regional discrimination. Impressively, the class CV predicted plot for XGBDA of the A-TEEM data (Fig. 2) indicated 100% correct classification of all varieties, with a distinct resolution among the different classes. Notably, it even separated the wine that was a 50:50 blend of Shiraz and Cabernet Sauvignon.

Data from the confusion matrix presented in Table 2 reiterates the 100% accuracy, with the maximum result achieved for all parameters for each varietal class with multi-block analysis. In comparison, analysis using only EEM data with XGBDA afforded somewhat lower accuracy (96.1% correct classification) and inferior model parameters, especially when sample numbers were low (i.e., Merlot and Shiraz/Cabernet Sauvignon, Table S5 of the Supplementary material). Fluorescence spectroscopy has been previously applied to white wine varietal authentication, achieving 100% correct classifications with PCA-LDA modelling for three Hungarian wine varieties (Sádecká and Jakubíková, 2020), and 97% correct classification for Chardonnay, Pinot Gris, Riesling and Sauvignon Blanc samples with soft independent modelling of class analogy (SIMCA) (Suciu et al., 2019). Therefore, the fluorescence technique can be accepted as a viable tool for varietal authentication of wine. Moreover, with the A-TEEM method, spectral data can be collected more rapidly in comparison to other spectrofluorometers. With the exceptional classification results, the underlying drivers of these classifications were further explored.

3.3. Identification of main fluorophores using PARAFAC

PARAFAC was conducted to extract the characteristic excitation and emission profiles of the main fluorophores arising from the studied wine samples. In PARAFAC, the number of components is selected considering the core consistency, where a high value (i.e., close to 100%) indicates the appropriateness of the model with only a few components, typically two to four with EEM data from organic matter (Murphy, Stedmon, Graeber, & Bro, 2013). However, that does not always apply from a chemical point of view, as it might not represent the samples with a complex chemical composition and containing many fluorophores (Airado-Rodríguez et al., 2011). Another powerful way of validating the PARAFAC model is to conduct split half analysis, where the data set is divided in half for a comparison of the similarity (Murphy et al., 2013). In this data set, the similarity measure of splits was 91.9% with five components (Fig. S2 of the Supplementary material), which indicated the stability of the model. EEM contour plots for the five component PARAFAC model are shown in Fig. 3.

The fluorescent compounds naturally present in wine and their fluorescent properties have been studied in the past with PARAFAC (Airado-Rodríguez et al., 2011; Airado-Rodríguez, Galeano-Díaz, Durán-Merás, & Wold, 2009). Therefore, λ_{ex} and λ_{em} values of components could be identified according to their maximum intensities (Fig. S3 of the Supplementary material). The first component (C1) had its maximum intensities at λ_{ex} 278 nm and λ_{em} 314 nm and could be tentatively assigned to catechin and related compounds. Component two (C2) appeared at λ_{ex} 286 nm and λ_{em} 352 nm and could be assumed to relate to anthocyanins. The third component (C3), appearing at a similar emission wavelength as component one (λ_{em} 316 nm) but with maximum excitation at λ_{ex} 286 nm, could be tentatively assigned to epicatechin and related compounds. Component four (C4) showed

Table 2
Confusion matrix results for XGBDA analysis of multi-block data (absorbance + EEM) for wines (in duplicate) according to variety.

Variety	N	Sensitivity %	Specificity %	Err %	Precision %	F1
Cabernet Sauvignon	142	100.00	100.00	0.00	100.00	1.00
Merlot	10	100.00	100.00	0.00	100.00	1.00
Shiraz	288	100.00	100.00	0.00	100.00	1.00
Shiraz/ Cabernet Sauvignon	2	100.00	100.00	0.00	100.00	1.00

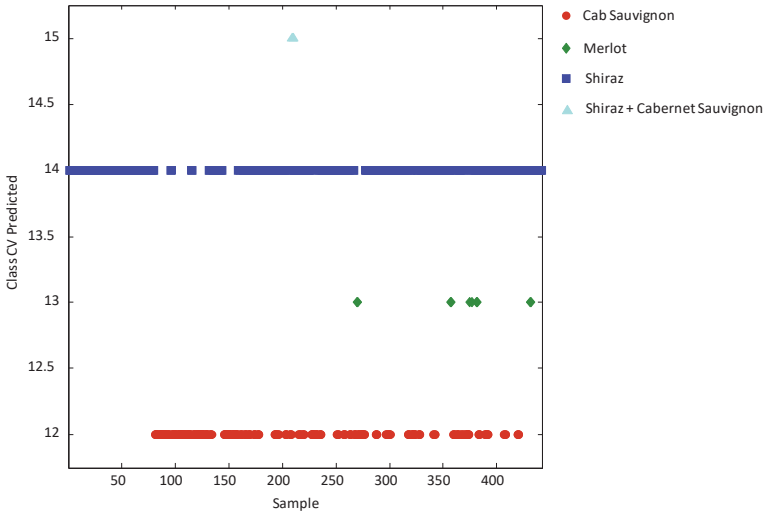


Fig. 2. Class CV predicted for variety from XGBDA analysis of A-TEEM data for red wines (n = 221, analysed in duplicate) from 10 Australian GIs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

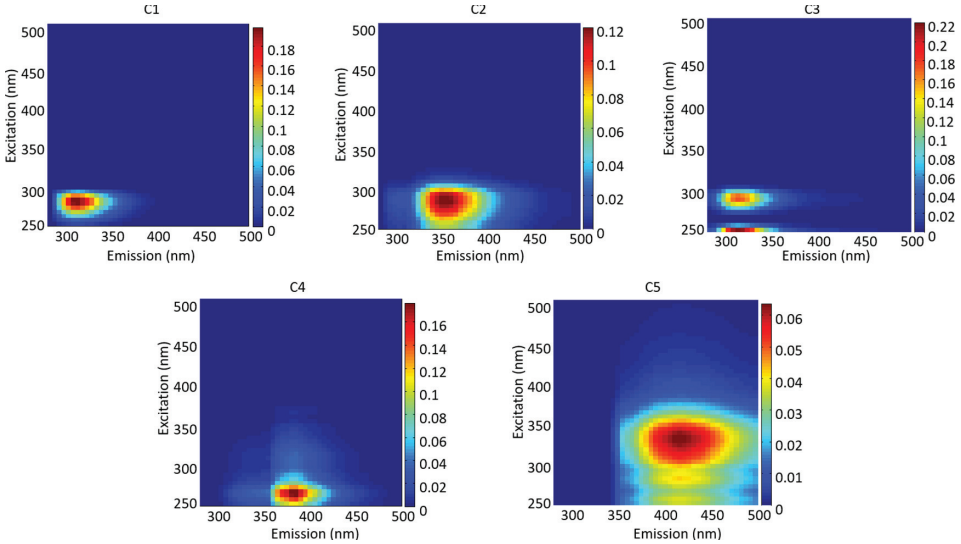


Fig. 3. EEM contour plots of the PARAFAC model with five components indicating the fluorescent properties of typical fluorophores in wine. Tentative assignments: C1, catechin; C 2, anthocyanin; C 3, epicatechin; C 4, phenolic acids; C 5, riboflavin).

maximum intensities at λ_{ex} 263 nm and λ_{em} 400 nm, which represents phenolic acids, such as gallic, vanillic, and ferulic acids. Component five (C5) has a broad signal with maximum intensities at around $\lambda_{ex}/\lambda_{em}$ 260–400/380–500 nm. Christensen, Nørgaard, Bro, and Engelsen (2006) tentatively assigned the spectral data to vitamins (riboflavin) in contrast to Schueuermann, Silcock, and Bremer (2018) who assigned them as phenolic compounds such as caffeic acid, *p*-coumaric acid or tyrosol. The present results compared well to the four component PARAFAC model of Airado-Rodríguez et al. (2009), where compounds such as catechin, epicatechin and phenolic acids are commonly revealed. However, in the current work it was also possible to identify a prominent component that could be tentatively assigned to anthocyanin. Ultimately, the PARAFAC components cannot be matched with a single compound but rather could agglomerate fluorescent molecules with related behaviour due to their functional groups. That is useful information from a sensory perspective (in as much as certain phenolics can relate to bitterness or mouthfeel, for example), although component clustering (data not shown) did not reveal anything about the samples according to region or variety in this case. This was most likely attributable to the data set consisting of unbalanced groups, causing overfitting of all the classes with the same loading. Nonetheless, further understanding could be derived from quantitative phenolic profiling of wines to identify specific drivers of variation among the samples (in terms of variety or region).

3.4. Prediction of concentrations of phenolic compounds with A-TEEM

Predicting analyte concentration from fluorescence spectra of an unknown sample can be achieved with multivariate calibration. A regression model can be built on fluorescence data of fluorophoric compounds versus their concentrations determined by a reference method and the model can be applied to quantifying the concentrations of fluorescent compounds in subsequent samples with simple and rapid fluorescence measurements (Sikorska et al., 2008). On that basis, phenolic concentration predictions were carried out in parallel to a study of Chilean wine (Gilmore et al., 2020), in which wines were analysed with A-TEEM and by HPLC as the reference method (which notably is much more time consuming than A-TEEM analysis). A range of phenolic compounds (n = 24) from various classes – including those tentatively identified in the PARAFAC modelling of the present work – were considered in that previous study and PLSR was applied to predict compound concentrations from A-TEEM data. Regression models were built for each compound separately as well as for totals of anthocyanins, flavan-3-ols, hydroxycinnamates, flavonols, and totals of low molecular weight phenolic (LMWP) compounds (sum of all flavan-3-ols, hydroxycinnamates, and flavonols) simultaneously (Gilmore et al., 2020). The same PLSR model was adopted in the current study to predict the concentration of anthocyanin glycosides (9), flavan-3-ols (2), flavonols and their glycosides (10), and hydroxycinnamates (3), as well as the five totals described above. A summary of PLSR statistics for the above-mentioned totals (which are indicative of the models for the individual compounds) appears in Table S6 of the Supplementary material, showing that strong correlations were achieved ($r > 0.99$) and the explanatory power of the model was high (adjusted $R^2 > 0.99$). Also, the SE was < 0.002 for the different variables and the low ratio of RMSECV/RMSEC ensured the robustness of the model (Gilmore & Chen, 2020). The potential benefit of spectroscopic methods (e.g., near-infrared) has been recognised for prediction of phenolic concentrations as it decreases analytical time and the cost of monitoring phenolic compounds (important for style and quality) in wines, specifically in commercial applications (Cozzolino et al., 2004). The present results illustrate that A-TEEM could also be used as a reliable method for phenolic quantification in addition to being an excellent approach to classification. Moreover, using fluorescence is advantageous over other spectroscopic techniques as it can measure compounds at low concentrations (Gilmore & Chen, 2020).

The phenolic concentrations obtained from the PLSR model were explored using one-way ANOVA, revealing significant differences among regions and varieties. In terms of variety, total anthocyanins were higher in monovarietal Cabernet Sauvignon wines in comparison to Shiraz and Merlot, which was mainly contributed by malvidin-3-*O*-glucoside (Table S7 of the Supplementary material). The anthocyanin profiles of grape skins can be considered for varietal characterisation due to the individual or total concentrations of anthocyanins varying among varieties (He et al., 2012). With regard to regional variations, each phenolic measure was highly significant ($p < 0.0001$) according to region when considering the 10 GIs (Table S8 of the Supplementary material). At a broad level, it was clear that total anthocyanins,

important to red wine colour, were low in Riverland samples in comparison to other regions, whereas higher amounts were present in wines from Western Australian regions (i.e., Frankland River and Margaret River, Table S8 of the Supplementary material). Biosynthesis of phenolics such as flavonols and anthocyanins is related to climatic characteristics of the regions such as light quality/quantity and temperature: specifically, lower temperatures increase the accumulation of anthocyanins in grapes and flavonol accumulation is enhanced with sunlight exposure (Blancquaert, Oberholster, Ricardo-da-Silva, & Deloire, 2019). As shown in Table S9 of the Supplementary material, Riverland has a higher mean January temperature (30.8 ℃) and growing degree days (GDD) of 2496 in comparison to Margaret River (19.4 ℃, 1712) and

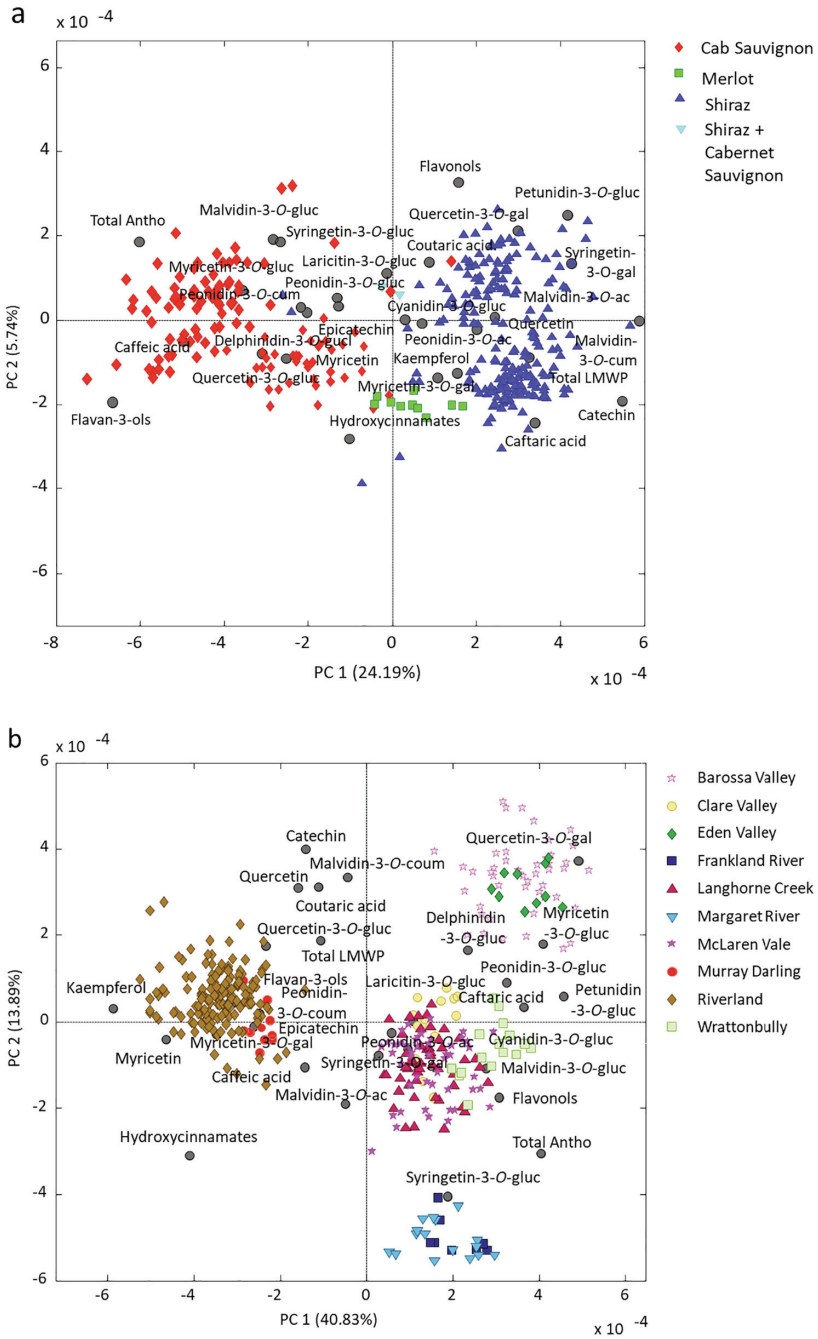


Fig. 4. PCA biplot showing scores for the wine samples according to a) variety and b) region of origin along with the loadings based on phenolic compound concentrations. Abbreviations: gluc, glucoside; gal, galactoside; ac, 3-*O*-acetylglucoside; coum, 3-*O*-coumaroylglucoside.

Frankland River regions (19.3 ℃, 1610). In their study with Australian Pinot Noir wines, Longo et al. (2020) also found that total anthocyanins were significantly different between regions, which might be attributable to the reaction of anthocyanins with tannin, with some regional differences in tannin extractability possibly being related to climate. Considering the multiple phenolic parameters being considered and the variation among samples, PCA of the data was undertaken to identify relationships and simplify the discussion.

3.4.1. PCA of quantitative phenolic data

The concentrations of phenolics were further explored with robust PCA with GLSW pre-processing and biplots were separately obtained for both region and variety. Despite the relatively low explained variance by the first two dimensions (30%) and lower still for other dimensions, clustering by variety can be observed for Shiraz and Cabernet Sauvignon along PC1, with Merlot and Shiraz/Cabernet Sauvignon blend separating along PC2 (Fig. 4a). Segregation of samples according to the variety could be ascribed to certain phenolic compounds. Compounds to the right along PC1, such as malvidin-3-*O*-coumaroylglucoside, syringetin-3-*O*-galactoside, petunidin-3-*O*-glucoside, and catechin were especially characteristic of Shiraz. On the left along PC1 were total anthocyanins, total flavan-3-ols and compounds such as caffeic acid, myricetin-3-*O*-glucoside, malvidin-3-*O*-glucoside, and peonidin-3-*O*-coumaroylglucoside, which were close to the Cabernet Sauvignon samples. Total hydroxycinnamates tended to characterise the Merlot samples in the lower half of the plot along PC2.

Considering the biplot according to region (Fig. 4b), 55% of the variance among the sample set can be explained by the first two PCs, and certain clustering by region can be observed. Given their relative proximity, it is interesting to see that Riverland and Murray Darling samples were clustered together on the left side of PC1 with a number of flavonols, particularly kaempferol and myricetin, and total flavan-3-ols. In the opposite direction were most of the anthocyanins, total anthocyanins, total flavonols, caftaric acid, and the other regions. PC2 tended to separate Barossa and Eden Valleys, and Margaret and Frankland Rivers from most of the other regions. The likely effect of terroir of the regions on phenolic composition of the wines can be seen in Fig. 4b, especially the separation of warmer climate regions such as Riverland and Murray Darling to the left. Clare Valley, Langhorne Creek, Wrattonbully and McLaren Vale samples were located relatively close together and nearer the centre of the plot, although there was slight separation of Clare Valley along PC2 and Wrattonbully along PC1.

In the same way as for the variety, a number of phenolic compounds can be related to those clusters around regions, such as syringetin-3-*O*-glucoside for the Western Australian cluster, and myricetin-3-*O*-glucoside, delphinidin-3-*O*-glucoside and quercetin-3-*O*-galctoside for the Barossa Valley and Eden Valley cluster. Also, cyanidin-3-*O*-glucoside, laricitin-3-*O*-glucoside, and peonidin-3-*O*-acetylglucoside were relevant to the McLaren Vale cluster, and epicatechin, myricetin-3-*O*-galactoside and peonidin-3-*O*-coumaroylglucoside were situated around the Riverland and Murray Darling cluster. It is worthwhile noting that the clusters also contained some wine blends from the same region. For instance, some samples belonging to Barossa Valley consisted of 95% Barossa Valley with 5% of wine from Eden Valley. Similarly, some of the McLaren Vale samples contained 90% McLaren Vale wine with 7% from Clare Valley and 3% from Langhorne Creek. Overall, the results showed that A-TEEM data encompasses sensitive information about samples that can be used for phenolic quantification via PLSR as well as correct classification of samples using XGBDA, making this a powerful platform for quality assessment (on the basis of phenolic profile) and authentication of wine.

4. Conclusion

A broad range of wines belonging to three varieties (Shiraz, Cabernet Sauvignon and Merlot) from 10 different Australian GIs were studied. Wines were evaluated after simple dilution by applying the A-TEEM technique in conjunction with XGBDA modelling, in order to differentiate the samples with respect to their variety and geographical origin. Excellent results were obtained in relation to cross-validation parameters (sensitivity, specificity, precision and F1 values) for each class. Highlighting the effectiveness of A-TEEM data with XGBDA in classifying the wines, an unrivalled accuracy of 100% was achieved for the varietal authentication and 99.7% for geographical discrimination. Furthermore, PARAFAC analysis tentatively revealed the most dominant fluorescent compounds in the set of wines that described the variability present in the EEMs. Robust calibration models developed using PLSR with HPLC reference data for prediction of phenolic concentrations were applied to the quantification of 24 relevant phenolic compounds and 5 compound class totals. PCA of the quantitative data helped reveal the differences according to variety and region and gave the first insight into the information embedded in the A-TEEM spectral data that underpins the classification models. Even though further research could be undertaken to strengthen the calibration models with more samples or to include more analytes, these results indicate that A-TEEM with chemometrics can be a valuable tool for the timely and novel quantification of individual phenolic compounds in wine, especially in an applied industry setting when compared to a more complex analytical method such as HPLC.

CRediT authorship contribution statement

Ranaweera K.R. Ranaweera: Conceptualization, Funding acquisition, Formal analysis, Investigation, Methodology, Visualization, Writing - original draft, Writing - review & editing. **Adam M. Gilmore:** Formal analysis, Investigation, Methodology, Writing - review & editing. **Dimitra L. Capone:** Conceptualization, Supervision, Writing - review & editing. **Susan E.P. Bastian:** Conceptualization, Funding acquisition, Supervision, Writing - review & editing. **David W. Jeffery:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

R.K.R.R., D.L.C., S.E.P.B. and D.W.J declare no conflict of interest. A. M.G., as an employee of HORIBA Instruments Inc., has no competing financial interests to declare. HORIBA Instruments Inc. did not influence the interpretation of results or decision to publish the manuscript.

Acknowledgements

We are immensely grateful to Warren Birchmore, Technical Viticulturist & Systems Manager on behalf of Accolade Wines, for the assistance with sample collection from the different wineries owned by Accolade Wines. Dr Andrew Jane, Quark Photonics Pty Ltd is thanked for facilitating Aqualog installation, training and service support. We are thankful to our UA colleagues, Claire Armstrong and Nick van Holst, and intern Jules Vallet from Institut Polytechnique LaSalle, for laboratory assistance. R.K.R.R. is appreciative of the financial support provided by an Australian Government Scholarship and Wine Australia supplementary scholarship (WA Ph1909). The Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au; project number IC170100008) is funded by the Australian Government with additional support from Wine Australia, Waite Research Institute, and industry partners. The University of Adelaide is a member of the Wine Innovation Cluster.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.

org/10.1016/j.foodchem.2021.130149.

References

Airado-Rodríguez, D., 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*(2), 257–264. https://doi.org/10.1016/j.jfca.2010.10.005.

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*(5), 1711–1720. https://doi.org/10.1021/jf8033623.

Aleixandre-Tudo, J. L., Buica, A., Nieuwoudt, H., Aleixandre, J. L., & du Toit, W. (2017). Spectrophotometric analysis of phenolic compounds in grapes and wines. *Journal of Agricultural and Food Chemistry, 65*(20), 4009–4026. https://doi.org/10.1021/acs.jafc.7b01724.

Azcarate, S. M., de Araújo Gomes, A., Alcaraz, M. R., Ugulino de Araújo, M. C., 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.

Blancquaert, E. H., Oberholster, A., Ricardo-da-Silva, J. M., & Deloire, A. J. (2019). Grape flavonoid evolution and composition under altered light and temperature conditions in Cabernet Sauvignon (*Vitis vinifera L.*). *Frontiers in Plant Science, 10*, Article 1062. https://doi.org/10.3389/fpls.2019.01062.

Cabrera-Bañegil, M., Hurtado-Sánchez, M. D. 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. https://doi.org/10.1016/j.foodchem.2016.09.152.

Chen, T., Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). USA: ACM. doi:10.1145/2939672.2939785.

Christensen, J., Nørgaard, L., Bro, R., & Engelsen, S. B. (2006). Multivariate autofluorescence of intact food systems. *Chemical Reviews, 106*(6), 1979–1994. https://doi.org/10.1021/cr050019q.

Coelho, C., Aron, A., Roullier-Gall, C., Gonsior, M., Schmitt-Kopplin, P., & Gougeon, R. D. (2015). Fluorescence fingerprinting of bottled white wines can reveal memories related to sulfur dioxide treatments of the must. *Analytical Chemistry, 87*(16), 8132–8137. https://doi.org/10.1021/acs.analchem.5b00388.

Cozzolino, D., Kwiatkowski, M. J., Parker, M., Cynkar, W. U., Dambergs, R. G., Gishen, M., et al. (2004). Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta, 513*(1), 73–80. https://doi.org/10.1016/j.aca.2003.08.066.

Dankowska, A. (2016). Advances in fluorescence emission spectroscopy for food authenticity testing. *In Advances in Food Authenticity Testing* (pp. 117–145). UK: Woodhead Publishing.

Dufour, É., Letort, A., Laguet, A., Lebecque, A., & Serra, J. N. (2006). Investigation of variety, typicality and vintage of French and German wines using front-face fluorescence spectroscopy. *Analytica Chimica Acta, 563*(1–2), 292–299. https://doi.org/10.1016/j.aca.2005.11.005.

Eigenvector Research. (2017). Confusionmatrix. Retrieved from: https://wiki.eigenvector.com/index.php?title=Confusionmatrix. Accessed 25.10.2020.

Ferrer-Gallego, R., Rodríguez-Pulido, F. J., Toci, A. T., & García-Estevez, I. (2020). Phenolic composition, quality and authenticity of grapes and wines by vibrational spectroscopy. *Food Reviews International, 1–29*. https://doi.org/10.1080/87559129.2020.1752231.

Gilmore, A. M. (2014). How to collect National Institute of Standards and Technology (NIST) traceable fluorescence excitation and emission spectra. In Y. Engelborghs, & A. J. W. G. Visser (Eds.), *1076. Fluorescence Spectroscopy and Microscopy (Methods and Protocols)* (pp. 3–27). Totowa, NJ: Humana Press.

Gilmore, A., Akaji, S., & Csatorday, K. (2017). Spectroscopic analysis of red wines with A-TEEM molecular fingerprinting. *Readout, E49*, 41–48. https://www.horiba.com/in/publications/readout/article/spectroscopic-analysis-of-red-wines-with-a-teem-molecular-fingerprinting-53035/.

Gilmore, A. M., & Chen, L. (2020). Water soluble fraction (WSF) contaminant detection using machine-learning absorbance-transmission excitation emission matrix (A-TEEM) spectroscopy. In I. Gannot (Ed.), *SPIE BiOS (11233, 112330H)*. San Francisco, CA: Society of Photo-Optical Instrumentation Engineers. https://doi.org/10.1117/12.2556434.

Gilmore, A. M., Schober, D., Penichet, I., Zincker, J., & González, A. (2020). Quality assessment of Cabernet Sauvignon from Chile based on simultaneous absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) analysis, *FACSS 47. Poster session presentation at SciX 2020 Virtual Conference.* Santa Fe, NM: FACSS Analytical Science and Innovation.

Godden, P., Wilkes, E., & Johnson, D. (2015). Trends in the composition of Australian wine 1984–2014. Australian Journal of Grape and Wine Research, 21(S1), 741-753. https://doi.org//10.1111/ajgw.12195.

He, F., Liang, N.-N., Mu, L., Pan, Q.-H., Wang, J., Reeves, M. J., et al. (2012). Anthocyanins and their variation in red wines I. Monomeric anthocyanins and their color expression. *Molecules, 17*(2), 1571–1601. https://doi.org/10.3390/molecules17021571.

Jiménez-Carvelo, A. M., González-Casado, A., Bagur-González, M. G., & Cuadros-Rodríguez, L. (2019). Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity – A review. *Food Research International, 122*, 25–39. https://doi.org/10.1016/j.foodres.2019.03.063.

Karoui, R., & Blecker, C. (2011). Fluorescence spectroscopy measurement for quality assessment of food systems—A review. *Food and Bioprocess Technology, 4*(3), 364–386. https://doi.org/10.1007/s11947-010-0370-0.

Kulmyrzaev, A. A., Karoui, R., De Baerdemaeker, J., & Dufour, E. (2007). Infrared and fluorescence spectroscopic techniques for the determination of nutritional constituents in foods. *International Journal of Food Properties, 10*(2), 299–320. https://doi.org/10.1080/10942910601045305.

Longo, R., Pearson, W., Merry, A., Solomon, M., Nicolotti, L., Westmore, H., et al. (2020). Preliminary study of Australian Pinot Noir wines by colour and volatile analyses, and the Pivot© profile method using wine professionals. *Foods, 9*(9), 1142. https://doi.org/10.3390/foods9091142.

Martin, A. E., Watling, R. J., & Lee, G. S. (2012). The multi-element determination and regional discrimination of Australian wines. *Food Chemistry, 133*(3), 1081–1089. https://doi.org/10.1016/j.foodchem.2012.02.013.

Murphy, K. R., Stedmon, C. A., Graeber, D., & Bro, R. (2013). Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Analytical Methods, 5*, 6557–6566. https://doi.org/10.1039/c3ay41160e.

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. *Food Chemistry, 335*, Article 127592. https://doi.org/10.1016/j.foodchem.2020.127592.

Ranaweera, R. K. R., Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). 3.33 - Authenticity and traceability in the wine industry: From analytical chemistry to consumer perceptions. In A. Cifuentes (Ed.), *3. Comprehensive Foodomics* (pp. 452–480). UK: Elsevier. https://www.sciencedirect.com/science/article/pii/B978008100596522876X.

Riovanto, R., Cynkar, W. U., Berzaghi, P., & Cozzolino, D. (2011). Discrimination between Shiraz wines from different Australian regions: The role of spectroscopy and chemometrics. *Journal of Agricultural and Food Chemistry, 59*(18), 10356–10360. https://doi.org/10.1021/jf202578f.

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 analysis. *Talanta, 153*, 278–284. https://doi.org/10.1016/j.talanta.2016.03.023.

Sádecká, J., & Jakubíková, M. (2020). Varietal classification of white wines by fluorescence spectroscopy. *Journal of Food Science and Technology, 57*(7), 2545–2553. https://doi.org/10.1007/s13197-020-04291-y.

Sádecká, J., Uríčková, V., Májek, P., & Jakubíková, M. (2019). Comparison of different fluorescence techniques in brandy classification by region of production. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 216*, 125–135. https://doi.org/10.1016/j.saa.2019.03.018.

Samokhvalov, A. (2020). Analysis of various solid samples by synchronous fluorescence spectroscopy and related methods: A review. *Talanta, 216*, Article 120944. https://doi.org/10.1016/j.talanta.2020.120944.

Schueuermann, C., Silcock, P., & Bremer, P. (2018). Front-face fluorescence spectroscopy in combination with parallel factor analysis for profiling of clonal and vineyard site differences in commercially produced Pinot Noir grape juices and wines. *Journal of Food Composition and Analysis, 66*, 30–38. https://doi.org/10.1016/j.jfca.2017.11.005.

Sikorska, E., Gliszczyńska-Świgło, A., Insińska-Rak, M., Khmelinskii, I., De Keukeleire, D., & Sikorski, M. (2008). Simultaneous analysis of riboflavin and aromatic amino acids in beer using fluorescence and multivariate calibration methods. *Analytica Chimica Acta, 613*(2), 207–217. https://doi.org/10.1016/j.aca.2008.02.063.

Suciu, R.-C., Zarbo, L., Guyon, F., & Magdas, D. A. (2019). Application of fluorescence spectroscopy using classical right angle technique in white wines classification. *Scientific Reports, 9*, Article 18250. https://doi.org/10.1038/s41598-019-54697-8.

Tan, J., Li, R., Jiang, Z.-T., Zhang, Y., Hou, Y.-M., Wang, Y.-R., et al. (2016). Geographical classification of Chinese Cabernet Sauvignon wines by data fusion of ultraviolet–visible and synchronous fluorescence spectroscopies: the combined use of multiple wavelength differences. *Australian Journal of Grape and Wine Research, 22*(3), 358–365. https://doi.org/10.1111/ajgw.12214.

Van Leeuwen, C., & Seguin, G. (2006). The concept of terroir in viticulture. *Journal of Wine Research, 17*(1), 1–10. https://doi.org/10.1080/09571260600633135.

Vergara, C., von Baer, D., Mardones, C., & Gutiérrez, L. (2011). Overview of chemical markers for varietal authentication of red wines. In S. E. Ebeler, G. R. Takeoka, & P. Winterhalter (Eds.), *1081. Progress in Authentication of Food and Wine* (pp. 101–111). Washington, DC: American Chemical Society.

Waterhouse, A. L. (2002). Wine phenolics. *Annals of the New York Academy of Sciences, 957*(1), 21–36. https://doi.org/10.1111/j.1749-6632.2002.tb02903.x.

Waterhouse, A. L., Sacks, G. L., & Jeffery, D. W. (2016). Introduction to phenolics. In A. L. Waterhouse, G. L. Sacks, & D. W. Jeffery (Eds.), *Understanding Wine Chemistry* (pp. 99–104). UK: John Wiley & Sons Ltd.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR

Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine

Ranaweera K. R. Ranaweera[a], Adam M. Gilmore[b], Dimitra L. Capone[a,c] Susan E. P. Bastian[a,c], David W. Jeffery[a,c,*]

[a] Department of Wine Science and Waite Research Institute, The University of Adelaide (UA), PMB 1, Glen Osmond, South Australia 5064, Australia

[b] HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, USA

[c] Australian Research Council Training Centre for Innovative Wine Production, UA, PMB 1, Glen Osmond, South Australia 5064, Australia

[*] Corresponding author.

E-mail address: david.jeffery@adelaide.edu.au (D.W. Jeffery).

Table of contents

Table S1. Summary of the different varieties and the regions (GIs) of the samples analysed in this study. . S-2
Table S2. LS mean and one-way ANOVA results of basic chemical measures for the different varieties ... S-2
Table S3. LS mean and one-way ANOVA results of chemical measures for the different regions S-2
Table S4. Confusion matrix results for EEM data of XGBDA analysis according to region of origin. S-3
Table S5. Confusion matrix results for EEM data of XGBDA analysis according to variety. S-3
Table S6. Summary of statistics obtained from PLSR of HPLC reference values with A-TEEM measures. S-3
Table S7. LS mean and one-way ANOVA results of concentrations of phenolics for the varieties S-4
Table S8. LS mean and one-way ANOVA results of concentrations of phenolics for the different regions. S-5
Table S9. Climatic data of region between 2018-2019 season S-6
Fig. **S1** Locations of the GIs for the samples collected from Australian regions S-6
Fig. S2 Split-half analysis for (**a**) excitation and (**b**) emission of the five components in the non-negativity constrained PARAFAC model. S-7
Fig. S3 Fluorescence loadings for (**a**) excitation and (**b**) emission of the five components in the non-negativity constrained PARAFAC model S-7

Table S1. Summary of the different varieties and the regions (GIs) of the samples analysed in this study.

Region	No. of samples from each variety				Total
	Shiraz	Cabernet Sauvignon	Merlot	Shiraz/Cabernet Sauvignon	
Barossa Valley	28	0	0	0	28
Clare Valley	8	0	0	0	8
Eden Valley	12	0	0	0	12
Frankland River	2	5	0	0	7
Langhorne Creek	15	10	0	0	25
Margaret River	0	17	0	0	17
McLaren Vale	13	12	0	0	25
Murray Darling	3	2	0	0	5
Riverland	59	20	5	0	84
Wrattonbully	4	5	0	1	10
Total	144	71	5	1	221

Table S2. Least squares means and one-way ANOVA results of basic chemical measures for the different varieties. Different letters within a column indicate statistically significant differences ($\alpha = 0.05$).

Variety	pH	Titratable Acidity (gL^{-1} of tartaric acid)	Alcohol % (v/v)
Shiraz	3.57 b	5.4 a	14.7 a
Cabernet Sauvignon	3.54 c	5.4 ab	14.4 c
Merlot	3.65 a	4.7 c	14.4 bc
Shiraz/Cabernet Sauvignon	3.53 c	5.3 b	14.6 b
p-value	< 0.0001	0.007	< 0.0001

Table S3. Least squares means and one-way ANOVA results of chemical measures for the different regions. Different letters within a column indicate statistically significant differences ($\alpha = 0.05$).

Region	pH	Titratable Acidity (gL^{-1} of tartaric acid)	Alcohol (% v/v)
Barossa Valley	3.57 ab	6.1 a	14.3 b
Clare Valley	3.55 bc	5.4 cd	14.9 a
Eden Valley	3.53 bc	5.7 bc	14.7 b
Frankland River	3.44 d	5.0 ef	14.1 b
Langhorne Creek	3.53 c	6.1 a	14.8 a
Margaret River	3.53 c	4.9 f	14.1 b
McLaren Vale	3.57 ab	6.0 ab	14.6 b
Murray Darling	3.51 cd	5.2 de	14.7 ab
Riverland	3.59 a	4.9 f	14.7 b
Wrattonbully	3.52 c	5.6 bc	14.4 bc
p-value	< 0.0001	< 0.0001	< 0.0001

Table S4. Confusion matrix results for EEM data of XGBDA analysis of wines (in duplicate) according to region of origin.

Region	N	Sensitivity%	Specificity%	Error%	Precision%	F1 Score
Barossa Valley	56	98.21	100.00	0.23	100.00	0.99
Clare Valley	16	100.00	99.77	0.23	94.12	0.97
Eden Valley	24	100.00	99.77	0.23	96.00	0.98
Frankland River	14	100.00	100.00	0.00	100.00	1.00
Langhorne Creek	50	100.00	100.00	0.00	100.00	1.00
Margaret River	34	100.00	100.00	0.00	100.00	1.00
McLaren Vale	50	98.00	100.00	0.23	100.00	0.99
Riverland	168	99.41	99.64	0.45	99.41	0.99
Wrattonbully	20	100.00	100.00	0.00	100.00	1.00
Murray Darling	10	100.00	100.00	0.00	100.00	1.00

Table S5. Confusion matrix results for EEM data of XGBDA analysis of wines (in duplicate) according to variety.

Variety	N	Sensitivity%	Specificity%	Err%	Precision%	F1
Cabernet Sauvignon	142	97.89	97.33	2.50	94.56	0.96
Merlot	10	30.00	99.77	1.80	75.00	0.42
Shiraz	288	97.91	94.80	3.17	97.24	0.98
Shiraz/Cabernet Sauvignon	2	50.00	100.00	0.23	100.00	0.67

Table S6. Summary of statistical results obtained from partial least squares regression of HPLC reference values (total per class of compound) with A-TEEM measures of wine samples.

	Correlation Coefficient (r)	R^2 (Adjusted)	Slope Value	Standard Error	RMSECV/ RMSEC
Flavan-3-ols	0.99909	0.99818	0.93692	0.00190	1.06
Hydroxycinnamates	0.99931	0.99862	0.99128	0.00175	1.35
Flavonols	0.99918	0.99836	0.95473	0.00184	1.20
LMWP	0.99986	0.99973	1.00176	0.00079	1.19
Anthocyanins	0.99955	0.99911	0.95739	0.00136	1.25

Table S7. Least squares means and one-way ANOVA results of concentrations of phenolic compounds (mgL^{-1}) for the different varieties. Different letters within a row indicate statistically significant differences ($\alpha = 0.05$).

	Cabernet Sauvignon	Shiraz	Merlot	Shiraz/Cabernet Sauvignon	p-value
Total anthocyanins	333.4 a	244.9 b	217.6 b	247.7 b	< 0.0001
Cyanidin-3-*O*-gluc[a]	1.3 a	0.7 c	0.7 c	1.0 b	< 0.0001
Delphinidin-3-*O*-gluc	14.1 a	8.5 c	7.6 c	9.9 b	< 0.0001
Malvidin-3-*O*-gluc	185.5 a	137.9 b	125.1 b	137.3 b	< 0.0001
Malvidin ac	70.9 a	54.9 b	45.8 c	45.9 c	< 0.0001
Malvidin coum	11.1 a	11.8 a	6.7 b	7.8 b	< 0.0001
Peonidin-3-*O*-gluc	4.3	4.3	2.5	3.5	0.094
Peonidin ac	2.7 a	2.2 a	1.6 b	1.9 b	< 0.0001
Peonidin coum	0.9 a	0.8 a	0.5 b	0.6 b	0.008
Petunidin-3-*O*-gluc	20.3 a	14.6 b	12.6 b	14.1 b	< 0.0001
Total LMWP	323.2	317.2	320.7	228.3	0.250
Total flavan-3-ols	98.5 a	88.2 a	122.4 a	52.9 b	< 0.0001
Catechin	42.9 a	43.7 a	64.5 a	25.4 b	< 0.0001
Epicatechin	49.3 a	39.7 a	50.9 a	28.0 b	< 0.0001
Total flavonols	147.4 a	152.2 a	116.3 b	124.8 b	0.035
Myricetin-3-*O*-gal	6.0 b	6.6 a	4.3 c	5.8 b	0.003
Myricetin-3-*O*-gluc	33.0 a	28.9 b	24.2 c	24.2 c	0.002
Myricetin	26.4 a	20.0 b	24.2 ab	12.3 c	< 0.0001
Quercetin-3-*O*-gal	24.3 b	28.3 a	17.9 c	25.7 b	< 0.0001
Quercetin-3-*O*-gluc	1.2 b	10.6 a	-0.4 b	1.3 b	< 0.0001
Quercetin	28.3	29.5	30.1	21.6	0.082
Laricitin-3-*O*-gluc	5.4 a	5.1 a	3.8 b	5.2 a	0.003
Syringetin-3-*O*-gal	3.5 b	5.6 a	3.4 bc	1.9 c	<0.0001
Syringetin-3-*O*-gluc	13.5 a	8.4 bc	6.6 c	12.6 ab	<0.0001
Kaempferol	2.9	2.9	2.3	2.1	0.701
Total hydroxycinnamates	81.9a	73.6 a	82.7 a	44.2 b	< 0.0001
Caftaric acid	60.3	58.1	60.6	39.3	0.068
Caffeic acid	9.7 a	7.4 b	9.9 a	4.8 c	< 0.0001
Coutaric acid	10.0 a	8.4 a	11.0 a	3.2 b	< 0.0001

[a]gluc, glucoside; gal, galactoside; ac, 3-*O*-acetylglucoside; coum, 3-*O*-coumaroylglucoside; LMWP, low molecular weight phenolic compounds

Table S8. Least squares means and one-way ANOVA results of concentrations of phenolic compounds (mgL^{-1}) for the different regions (all significantly different, $p < 0.0001$). Different letters within a row indicate statistically significant differences ($\alpha = 0.05$).

	Eden Valley	Frankland River	Barossa Valley	Margaret River	Clare Valley	McLaren Vale	Murray Darling	Langhorne Creek	Riverland	Wrattonbully
Total anthocyanins	348.3 b	430.6 a	311.3 bc	429.3 a	319.5 bc	305.1 bc	221.5 de	272.5 cd	194.6 e	258.0 cd
Cyanidin-3-*O*-gluc[a]	1.4 ab	1.8 a	1.2 bc	1.8 a	1.2 bc	1.1 bc	0.7 de	1.0 cd	0.4 e	1.0 d
Delphinidin-3-*O*-gluc	18.2 ab	17.0 abc	15.0 bc	18.9 a	12.9 cd	11.2 cd	8.3 de	9.7 d	5.0 e	9.8 d
Malvidin-3-*O*-gluc	192.2 b	235.0 a	170.9 bc	238.2 a	173.8 bc	168.2 bc	127.1 de	152.0 cd	113.5 e	142.1 d
Malvidin ac	74.9 b	99.9 a	68.2 b	94.5 a	71.4 b	67.3 b	43.2 d	56.3 c	44.4 d	49.3 cd
Malvidin coum	16.6 a	16.3 a	16.1 a	14.9 a	14.6 a	12.6 b	8.8 d	10.6 c	8.2 d	8.1 d
Peonidin-3-*O*-gluc	8.2 a	4.8 bc	7.5 a	5.6 b	5.1 b	4.1 c	3.2 c	3.5 c	2.6 d	3.1 c
Peonidin ac	3.9 a	3.5 ab	3.4 ab	3.6 ab	3.0 bc	2.5 cd	2.0 ef	2.2 de	1.5 f	1.7 f
Peonidin coum	1.5 a	1.4 a	1.4 a	1.3 a	1.2 ab	1.0 b	0.5 d	0.8 c	0.5 d	0.5 d
Petunidin-3-*O*-gluc	23.7 ab	24.8 ab	21.4 b	25.7 a	20.3 b	18.2 bc	14.1 de	15.6 cd	10.6 d	14.6 e
Total LMWP	409.2 a	360.1 ab	379.3 a	337.5 b	374.7 ab	331.5 b	297.2 bc	269.9 c	295.9 c	223.8 d
Total flavan-3-ols	118.5 a	91.7 bc	98.6 abc	92.3 bc	96.9 abc	85.9 c	92.1 bc	65.7 d	99.8 ab	53.4 e
Catechin	55.6 a	35.8 c	47.3 ab	36.5 c	39.1 bc	36.0 c	41.2 abc	27.7 c	52.5 a	28.9 c
Epicatechin	53.1 a	51.9 a	44.9 b	50.9 a	49.2 ab	44.6 b	42.6	35.3 c	41.4 b	28.5 d
Total flavonols	193.4 ab	174.4bc	200.8 a	153.8 c	198.1 ab	174.6 b	125.1 de	139.3 cd	119.3 e	120.9 e
Myricetin-3-*O*-gal	9.3 a	6.5 cd	9.7 a	6.1 d	8.9 ab	7.6 bc	5.3 de	6.1 d	4.4 e	5.6 d
Myricetin 3-*O*-gluc	44.0 a	33.1 bc	42.8 a	33.5 b	42.4 ab	36.8 ab	26.4 c	27.5 c	21.5 c	23.2 c
Myricetin	23.4 b	29.8 ab	21.4 b	31.5 a	23.2 b	21.7 b	25.3 b	18.7 b	21.5 b	13.0 c
Quercetin-3-*O*-gal	35.3 ab	29.2 bc	39.1 a	23.5 c	37.0 ab	32.3 b	19.9 d	26.4 c	20.0 d	24.9 c
Quercetin-3-*O*-gluc	14.4 ab	1.2 cd	19.2 a	1.3 d	14.1 ab	10.6 b	0.5 cd	3.1 c	5.3 c	1.1 cd
Quercetin	35.0 a	34.8 a	33.3 a	28.1 bc	35.1 a	29.6 b	28.7 bc	26.6 c	27.8 bc	19.5 d
Laricitin-3-*O*-gluc	5.9 a	6.6 a	6.6 a	5.8 a	6.8 a	6.3 a	4.4 cd	5.3 b	3.9 d	5.1 bc
Syringetin-3-*O*-gal	6.9 ab	4.6 bc	8.0 a	3.2 c	7.0 ab	5.5 b	3.1 c	3.4 c	4.4 c	1.7 c
Syringetin-3-*O*-gluc	6.0 f	19.8 a	8.4 e	16.5 ab	12.3 d	14.1 c	6.1 f	12.1 d	6.6 f	15.0 bc
Kaempferol	4.6 a	2.1 de	4.3 a	2.6 c	3.4 b	2.9 bc	3.1 bc	2.2 d	2.7 c	1.5 e
Total hydroxycinnamates	93.1 ab	102.1 a	73.0 bc	101.5 a	75.8 bc	70.3 bc	80.4 ab	62.6 c	77.1 b	44.4 d
Caftaric acid	75.6 a	77.4 a	60.6 b	74.8 a	60.8 ab	55.5 b	57.7 b	49.7 b	56.8 b	38.4 c
Caffeic acid	9.1 b	9.8 ab	7.4 bc	10.4 a	7.9 b	7.8 b	9.3 ab	6.6 c	8.6 b	4.7 d
Coutaric acid	10.6 ab	9.4 bc	9.0 bc	11.5 a	8.3 bc	7.4 c	12.7 a	6.5 c	9.9 b	2.5 d

[a]gluc, glucoside; gal, galactoside; ac, 3-*O*-acetylglucoside; coum, 3-*O*-coumaroylglucoside; LMWP, low molecular weight phenolic compounds

Table S9. Climatic data of the regions for season 2018-2019, sourced from Wine Australia Marketing Insights (https://www.wineaustralia.com/market-insights/regional-snapshots).

Region	Mean-January Temperature (°C)	Annual Rainfall (mm)	Growing Degree Days (GDD)
Barossa Valley	24.8	366	2120
Clare Valley	25.5	391	2156
Eden Valley	24.4	347	2031
Frankland River	23.3	614	1610
Langhorne Creek	23.4	316	2087
Margaret River	19.4	898	1712
McLaren Vale	23.3	590	1989
Murray Darling	27.3	269	2845
Riverland	30.8	149	2496
Wrattonbully	21.9	496	1688

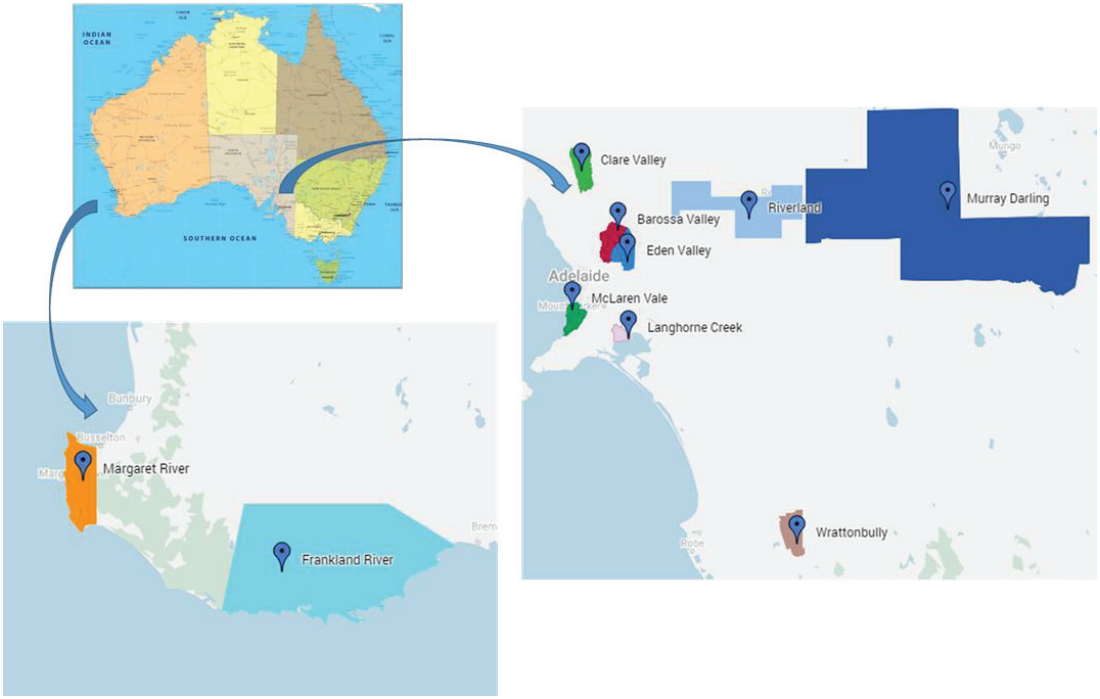


Fig. S1 Locations of the GIs for the samples collected from Australian regions: Margaret River and Frankland River from Western Australia; Clare Valley, Barossa Valley, Eden Valley, Riverland, McLaren Vale, Langhorne Creek, and Wrantonbully from South Australia; and Murray Darling contained within the zones of New South Wales and North West Victoria.

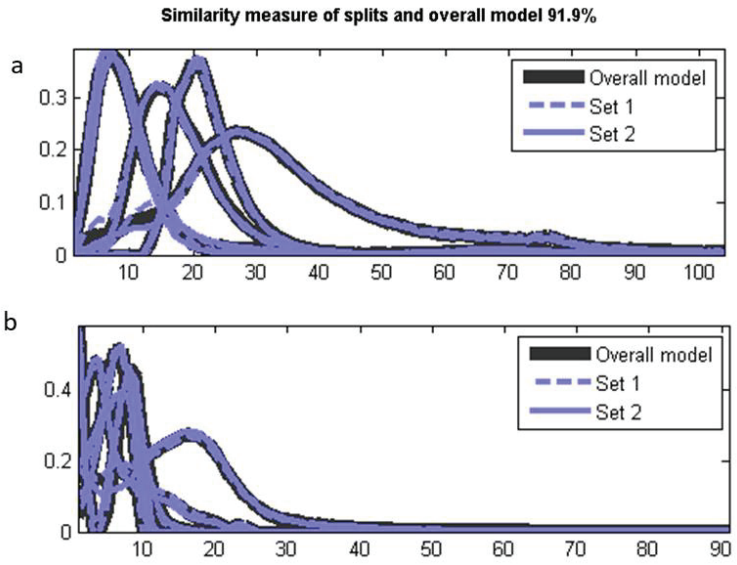


Fig. S2 Split-half analysis for (**a**) excitation and (**b**) emission of the five components in the non-negativity constrained PARAFAC model.

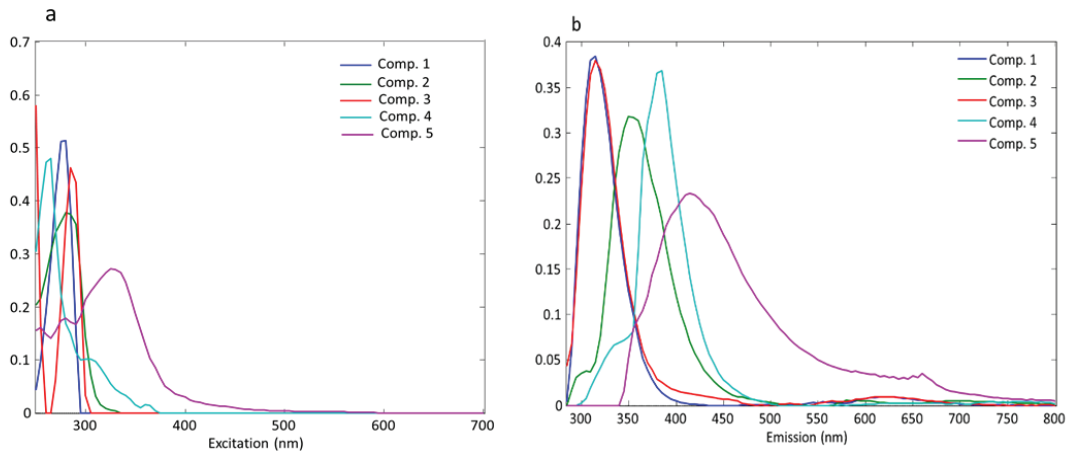


Fig. S3 Fluorescence loadings for (**a**) excitation and (**b**) emission of the five components in the non-negativity constrained PARAFAC model. Tentative assignments: Comp. 1, catechin; Comp. 2, anthocyanin; Comp. 3, epicatechin; Comp. 4, phenolic acids; Comp. 5: riboflavin.

Chapter 5

Manuscript for publication

Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine

Ranaweera, R. K. R., Bastian, S. E. P., Gilmore, A. M., Capone, D. L., & Jeffery, D. W. (2021). Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine. Food Chemistry, (Submitted for publication).

Statement of Authorship

Title of Paper	Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Ranaweera, R. K. R., Bastian, S. E. P., Gilmore, A. M., Capone, D. L., & Jeffery, D. W. (2021). Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine. Food Chemistry

Principal Author

Name of Principal Author (Candidate)	Ranaweera K. R. Ranaweera		
Contribution to the Paper	Contributed to the experimental design. Performed A-TEEM analysis for the wine samples. Carried out data collection, processing, and applied statistical techniques, created models and interpreted data for the manuscript. Drafted the manuscript, and reviewed and edited the final draft for publication.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	12/10/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Susan E.P. Bastian		
Contribution to the Paper	Contributed to the planning and design of experiments. Coordinated the wine sample acquisition. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	15/11/2021

Name of Co-Author	Adam M. Gilmore		
Contribution to the Paper	Contributed to design of methodology for wine analysis. Applied statistical techniques in formal analysis of the study data. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	2021/11/15

Name of Co-Author	Dimitra L. Capone		
Contribution to the Paper	Contributed to the planning and design of experiments. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	15/11/2021

Name of Co-Author	David W. Jeffery		
Contribution to the Paper	Original conceptualisation and design of the study. Acquired funding and provided resources for the project leading to this publication. Supervised the project and undertook project administration throughout the study. Critically evaluated, reviewed, edited and revised the manuscript for publication. Acted as the corresponding author at all stages.		
Signature		Date	12/10/2021

Food Chemistry

Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine

--Manuscript Draft--

Manuscript Number:	FOODCHEM-D-21-09175
Article Type:	Research Article (max 7,500 words)
Keywords:	Extreme gradient boosting; chemometrics; terroir; regionality; subregion; authentication; provenance
Corresponding Author:	David W Jeffery, PhD The University of Adelaide - Waite Campus Glen Osmond, AUSTRALIA
First Author:	Ranaweera K. R. Ranaweera
Order of Authors:	Ranaweera K. R. Ranaweera
	Susan E. P. Bastian
	Adam M. Gilmore
	Dimitra L. Capone
	David W Jeffery, PhD
Abstract:	Authentication of wine can be considered at different scales, with classification according to country, province/state, or appellation/wine producing region. An absorbance-transmission and excitation-emission matrix (A-TEEM) technique was applied for the first time to examine intraregional differences, using Shiraz wines (n = 186) produced during three vintages from five subregions of Barossa Valley and from Eden Valley. Absorption spectra and EEM fingerprints were modelled as a multi-block data set for initial exploration with k-means cluster analysis and principal component analysis, and then with machine learning modelling using extreme gradient boosting discriminant analysis (XGBDA). Whereas some clustering was evident with the initial unsupervised approaches, classification with XGBDA afforded an impressive 100% correct class assignment for subregion and vintage year. Extending the utility and novelty of the A-TEEM approach, predictive models for chemical parameters (alcohol, glucose + fructose, pH, titratable acidity, and volatile acidity) were also validated using A-TEEM data with XGB regression.
Suggested Reviewers:	Stamatina Kallithraka Agricultural University of Athens Department of Food Science and Human Nutrition: Geoponiko Panepistemio Athenon Tmema Epistemes Trophimon kai Diatrophes tou Anthropou stamatina@aua.gr Expertise in analytical and wine chemistry, wine chemical composition, chemometrics, and authenticity
	Hélène Nieuwoudt Stellenbosch University hhn@sun.ac.za Expertise in spectroscopy, wine analysis, phenolic compounds, chemometrics, and authenticity
	António Jordão Polytechnic Institute of Viseu: Instituto Politecnico de Viseu antoniojordao@esav.ipv.pt Expertise in wine chemistry, phenolic compounds, and wine technology
	Andrea Versari University of Bologna: Universita di Bologna andrea.versari@unibo.it Expertise in chemical and sensory analysis of wine, spectroscopy, and chemometrics

	Antonio Morata Universidad Politecnica de Madrid antonio.morata@upm.es Expertise in wine production technology, wine analysis, pigments and phenolic compounds

Highlights

- Shiraz wine subregion and vintage modelled with multi-block analysis of A-TEEM data
- Classification by unsupervised and supervised machine learning techniques explored
- Fingerprints from A-TEEM may be applied to the verification of subregional terroirs
- A-TEEM can be used to simultaneously predict the basic chemical parameters of wine

Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine

Ranaweera K. R. Ranaweera[a], Susan E. P. Bastian[a,b], Adam M. Gilmore[c], Dimitra L. Capone[a,b], David W. Jeffery[a,b,*]

[a] Department of Wine Science and Waite Research Institute, The University of Adelaide (UA), PMB 1, Glen Osmond, South Australia 5064, Australia

[b] Australian Research Council Training Centre for Innovative Wine Production, UA, PMB 1, Glen Osmond, South Australia 5064, Australia

[c] HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, USA

[*] Corresponding author.

E-mail address: david.jeffery@adelaide.edu.au (D.W. Jeffery).

Running title:

A-TEEM with machine learning for intraregional classification of Barossa Shiraz

Abstract

Authentication of wine can be considered at different scales, with classification according to country, province/state, or appellation/wine producing region. An absorbance-transmission and excitation-emission matrix (A-TEEM) technique was applied for the first time to examine intraregional differences, using Shiraz wines (n = 186) produced during three vintages from five subregions of Barossa Valley and from Eden Valley. Absorption spectra and EEM fingerprints were modelled as a multi-block data set for initial exploration with k-means cluster analysis and principal component analysis, and then with machine learning modelling using extreme gradient boosting discriminant analysis (XGBDA). Whereas some clustering was evident with the initial unsupervised approaches, classification with XGBDA afforded an impressive 100% correct class assignment for subregion and vintage year. Extending the utility and novelty of the A-TEEM approach, predictive models for chemical parameters (alcohol, glucose + fructose, pH, titratable acidity, and volatile acidity) were also validated using A-TEEM data with XGB regression.

Keywords:

Extreme gradient boosting, chemometrics, terroir, regionality, subregion, authentication, provenance

1. Introduction[1]

The ability to authenticate wines of provenance is an important consideration for the global wine industry, not only to prevent fraud but also to enhance consumer confidence and create value in the products being purchased. This can be bound with the concept of terroir, which relates to location-specific interactions – among topography, climate, soil, viticultural practices, and winemaking traditions – that influence production of different grape varieties and wines (Sáenz-Navajas & Jeffery, 2021; Souza Gonzaga et al., 2021). Considering the importance region and terroir, regulatory measures such as protected designation of origin (PDO) or protected geographical indication (PGI) have been introduced to preserve the authenticity, quality and typicity of wine from defined regions. In addition, fine wine regions have often gained their reputation for producing wines from well-suited grape varieties, such as Cabernet Sauvignon and Merlot blends from Bordeaux, Chardonnay from Napa Valley, Tempranillo from Rioja, Sauvignon Blanc from Marlborough, and Shiraz from Barossa Valley (Sáenz-Navajas & Jeffery, 2021).

Aside from being used to define wine regions, the distinctiveness of different terroirs has also been recognised for vineyard sites, such as in Mosel and Burgundy, as well as for wine estates of Bordeaux (Bastian & Iland OAM, 2019). Despite not having the depth of wine production history of Europe, Australia understands the value that can be associated with unique terroirs, with interest in capitalising on Australian terroirs extending to 'subregionalisation', which could enhance the sense of place upon linking wine distinctiveness to terroir (Bramley & Ouzman, 2021). Indeed, several wine regions in Australia have been identified by Wine Australia (2021) as subregions within the established Geographical Indication (GI) regions, such as High Eden in the Eden Valley GI and Frankland River in Great Southern GI, based on their distinctive terroir that influences wine characteristics. Continued research that verifies the connection between distinctive wines and unique

[1] Abbreviations: A-TEEM, absorbance-transmission and fluorescence excitation emission matrix; CA, cluster analysis; DA, discriminant analysis; GI, Geographical Indication; ICP-MS, inductively coupled plasma-mass spectrometry; LDA, linear discriminant analysis; PCA, principal component analysis; PLSDA, partial least squares discriminant analysis; SVM, support vector machine; PDO, protected designation of origin; PGI, protected geographical indication; XGB, extreme gradient boosting; XGBDA, extreme gradient boosting discriminant analysis.

terroirs is required, however, to provide scientific evidence that could support intraregional terroir zoning among wines of a given grape variety, as in a data-driven approach mentioned by Bramley and Ouzman (2021).

Highlighting some Australian research endeavours, projects have been established to understand the subregional terroirs in South Australian GIs. The Barossa Grounds project was based on Robinson and Sandercock (2014) having distinguished five subregions in the Barossa Valley: Northern Grounds, Central Grounds, and Southern Grounds, along with Eastern Edge and Western Ridge. Another project in McLaren Vale (Scarce Earth program) recognised nineteen different wine districts (Bekkers, 2012), and in Clare Valley, Werner and Roche (2016) defined five subregions (Rocks project). These projects were mainly based on climatic and topographic indices and soil studies, although other examples include research involving sensory analysis (Johnson et al., 2013; Kustos et al., 2020; Pearson et al., 2021; Souza Gonzaga et al., 2021) soil microbiology (Zhou et al., 2021), grape and wine chemistry (Chen et al., 2019), and grapevine epigenetics (Xie et al., 2017), conducted in an attempt to understand the influence of terroir on subregional variations. Beyond these studies, the question remained whether variations ascribed to intraregional differences in terroir could be demonstrated through chemical analysis of wine for the purpose of authentication, especially considering the challenge of authenticating products that are produced in close proximity.

Distinguishing wines at a subregional level has been approached by studying elements, which may be related back to vineyard soil. In their research, Coetzee et al. (2014) classified wines according to provenance from a single wine region in South Africa based on different estates, using inductively coupled plasma-mass spectrometry (ICP-MS). They identified several elements (B, Ba, Cs, Cu, Mg, Rb, Sr, Tl and Zn) as suitable indicators of wine origin and achieved about 80% correct classification with a combination of cluster analysis (CA) and discriminant analysis (DA). Also applying ICP-MS (and ICP-optical emission spectroscopy), a preliminary study by Aceto et al. (2020) aimed to authenticate Barbera d'Asti and Nizza wines from two geographically overlapping zones in Piedmont,

Italy, but concluded that the elemental profile of wine reflected oenological practices rather than soil variations, due to the proximity of the production zones.

Although accurate classification with ICP-MS can be achieved when considering diverse GI (Martin et al., 2012; Ranaweera, Gilmore, et al., 2021a), difficulty can be encountered when discriminating wines from GI within a state (e.g., South Australia), indicating that this technique may lack the necessary sensitivity (Martin et al., 2012). In contrast, the improved accuracy afforded when using UV-Vis absorbance and fluorescence data from an absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) approach in combination with machine learning classification offered the prospect of segregating wines from adjacent regions such as Clare, Barossa and Eden Valleys with 99.7 % accuracy, according to wine molecular fingerprints (Ranaweera, Gilmore, et al., 2021b). Furthermore, the speed and ease of use of a spectral approach was appealing, with UV-Vis and fluorescence spectroscopy offering particular advantages in terms of sensitivity, specificity, and accuracy (Ranaweera, Capone, et al., 2021).

To identify patterns among the samples and to condense the variables associated with spectroscopic data, exploratory data analysis is undertaken (Cozzolino et al., 2009) with unsupervised chemometric methods such as principal component analysis (PCA) and CA. For classification, however, supervised multivariate methods are often applied for wine authentication purposes, including partial least squares discriminant analysis (PLSDA), linear discriminant analysis (LDA), and support vector machine (SVM) (Ranaweera, Capone, et al., 2021). Additionally, an extreme gradient boosting (XGB) machine learning technique was recently applied to wine for the first time, achieving 100% accuracy in the classification of regional Cabernet Sauvignon wines using EEM data (Ranaweera, Gilmore, et al., 2021a). In contrast to classification, calibration models need to be generated for quantitative analysis, with PLS regression and principal component regression generally employed to avoid issues with noise and correlations in the data (Cozzolino et al., 2009).

Following the previous work (Ranaweera, Gilmore, et al., 2021a, 2021b), this study aimed to test whether A-TEEM could be used discriminate wine at a subregional level for the first time based

on the hypothesis that differences according to terroir would be chemically evident. Shiraz wines from Eden Valley GI and from five different subregions within the Barossa Valley GI of South Australia produced in three consecutive vintages were analysed by A-TEEM for classification according to origin and year of production. Unsupervised chemometric analyses with PCA and k-means clustering were undertaken as well as a classification modelling with extreme gradient boosting discriminant analysis (XGBDA). Additionally, A-TEEM data were modelled against reference methods in a novel approach to predicting basic wine chemical parameters using XGB regression.

2. Material and methods

2.1 Chemicals

HPLC gradient grade absolute ethanol and analytical grade 37% hydrochloric acid (HCl) were purchased from Chem-Supply (Port Adelaide, SA, Australia). High purity water was obtained from a Milli-Q purification system (Millipore, North Ryde, NSW, Australia).

2.2 Wine samples

A total of 186 samples of experimental or unreleased commercially produced Shiraz wines from five subregions of Barossa Valley[2] (Northern Grounds, Central Grounds, Southern Grounds, Western Ridge, Eastern Edge) and Eden Valley, South Australia, from 2018, 2019 and 2020 vintages were obtained (Table A.1 and Fig. A.1 of Appendix A).

2.3 Analytical procedures for basic chemical parameters

Wine pH and titratable acidity (TA) measurements were obtained with a Mettler Toledo T50 autotitrator, and alcohol content (percentage by volume) was measured with an Anton Paar Alcolyser. Volatile acidity (VA) as acetic acid and glucose/fructose values were determined enzymatically by Commercial Services at the Australian Wine Research Institute using a discrete analyser.

[2] Note that these subregions were not officially recognised in the Australian GI system. For simplicity, the term subregion is often used even if it relates to Eden Valley, which is defined as a region.

2.4 Analytical procedure for A-TEEM

Wine samples were analysed in duplicate by the A-TEEM technique according to Ranaweera, Gilmore, et al. (2021b). Samples were centrifuged at 9300 × g for 10 min and diluted with 50% aqueous ethanol that had been adjusted to pH 2 with HCl and degassed by vacuum filtration (0.45 µm PTFE membrane). Considering Beer-Lambert law, the wine-to-solvent dilution factor was determined to be 1:150 (Gilmore, 2014). After dilution, samples were mixed thoroughly with a benchtop vortex and sonicated to remove air bubbles. Samples were analysed with a HORIBA Scientific Aqualog spectrophotometer (version 4.2, Quark Photonics, Adelaide, SA, Australia) using a Hellma type 1FL (1 cm path length) Macro Fluorescence cuvette (Sigma-Aldrich, Castle Hill, NSW, Australia) with the same instrument settings as reported by Ranaweera, Gilmore, et al. (2021b) (i.e., excitation wavelength range of 240–800 nm with a 5 nm increment under medium gain and 0.2 s integration time; emission wavelength range of 242–824 nm with a 4.66 nm increment as set by the instrument). Absorbance spectra (240–700 nm) and EEMs were recorded with data acquisition undertaken using Origin software (version 8.6, OriginLab Corporation, Massachusetts, USA). Pre-processing of EEM data involved normalisation according to the water Raman scattering units for the specified emission conditions, and correcting for the influence of inner filter effects, solvent background, dark detector signals, and Rayleigh masking to eliminate spectral distortion (Gilmore et al., 2017).

2.5 Statistical analyses

Multi-block analysis was applied for A-TEEM data as described previously (Ranaweera, Gilmore, et al., 2021b), with EEM data unfolded into a two-way array using transform unfold multiway (mode 1) and absorbance data combined using the multi-block tool in Solo software (version 8.7.1, Eigenvector Research, Inc., Manson, WA, USA). Absorbance variables were baseline corrected and the multi-block data were adjusted for dilution factor and integration time. For unsupervised pattern recognition, cluster analysis with partitional k-means (k = 3) was applied for the full multi-block data set, and multi-block variable M2V (combined value of EEM and absorbance) was selected in Solo+MIA software (version 8.9.2). PCA with singular value decomposition and

autoscale pre-processing with seven principal components was also carried out for single vintages to discover subregional patterns using Solo software (version 8.7.1). Samples were classified according to subregional origin and vintage year using XGBDA after PLS compression using 10 latent variables, with mean centring pre-processing and decluttering with generalised least squares weighting at 0.2 to both calibrate and cross-validate (k =10, Venetian blinds procedure). For the evaluation of the model, confusion matrix score probabilities were considered according to previous studies (Ranaweera, Gilmore, et al., 2021a, 2021b). Basic chemical parameters were analysed by one-way analysis of variance (ANOVA) and Tukey's honestly significant difference (HSD) pairwise comparisons with a confidence interval of 95% ($\alpha = 0.05$) using XLSTAT (version 2019.03.02, Addinsoft, Boston, USA). For prediction of basic chemical parameters, calibration models were constructed with the XGB regression algorithm (Solo software, version 8.7.1) using A-TEEM data and chemical data obtained using reference methods outlined in section 2.3. Root mean square error of calibration (RMSEC), root mean square error of cross-validation (RMSECV) (Venetian blinds with 10 splits), and coefficients of determination for calibration and cross-validation (R^2 cal, R^2 CV) were used to measure the effectiveness of the correlations.

3. Results and Discussion

3.1 Exploratory data analysis

Building on previous encouraging results using A-TEEM for authentication of geographical origin of Australian wines from different regions (Ranaweera, Gilmore, et al., 2021a, 2021b), the current study sought to challenge the method even further, concentrating on discrimination at a subregional level for wines prepared in three consecutive vintages. This innovative approach addressed the need for data-driven methods (in this case based on chemistry) that help verify the effect of different terroirs on wine distinctiveness.

Considering the complexity of the multi-block data set, cluster analysis was first conducted as an unsupervised technique to identify patterns in the data. K-means partitional clustering showed clear groupings of wine samples according to the year of vintage (Fig. 1). Vintage was deemed to

have a strong effect on chemical composition that overshadowed any influence of the subregions, which is consistent with other work (e.g., Niimi et al., 2021; Niimi et al., 2020; Roullier-Gall et al., 2014). This could be related to the climatic variations between seasons (Lorrain et al., 2011) in terms of temperature, annual rainfall, and growing degree days, as shown in Table A.2 of Appendix A, which translated into discernible changes in the A-TEEM data according to underlying chemical profiles of the wines.

As subregional variations were not evident according to cluster analysis (data not shown), PCA was applied separately for each vintage to further explore whether unsupervised analysis of A-TEEM data could separate the subregions. Fig. 2 shows the distribution of samples by subregion according to the first three principal components, which explained a modest combined total of 36%, 24%, and 17% of variation for 2018, 2019, and 2020, respectively. Quite impressively, Fig. 2a shows rather tight groupings for all subregions in 2018, whereas samples from 2019 (Fig. 2b) and 2020 (Fig. 2c) were not so prominently differentiated. Nonetheless, separation of EV and NG was still reasonably apparent, along with a tendency for overlap between SG and CG. As shown in Table A.2 of Appendix A, 2020 was a relatively cool year for the Barossa Valley and more similar to that expected for Eden Valley, whereas 2019 had relatively low rainfall in both regions and was uncharacteristically warm for Eden Valley. Barossa Valley also experienced higher annual rainfall than Eden Valley in 2019 and 2020. Such factors likely resulted in less discernible differences among the wines for 2019 and especially for 2020, based on the well-known effect of vintage on grape composition as a result of variations in climate (van Leeuwen et al., 2004).

The PCA results in Fig. 2 imply an influence of subregion proximity at a basic level, as elucidated from Fig. A.1 of Appendix A, where EV is more remote from other subregions, although there would be other underlying factors. Using a data-driven approach with soil parameters such as available water capacity and carbon exchange capacity, along with climatic data including growing season rainfall, mean January temperature, and growing degree days, Bramley et al. (2020) also differentiated Barossa Valley and Eden Valley GIs, but not subregions of the Barossa Valley. In

192 contrast, using spectroscopic methods and chemometrics more akin to the current approach, Riovanto
193 et al. (2011) observed no clear differentiation of Shiraz wines from GI within South Australia, which
194 overlapped in the PCA score plots whether they were in close proximity or not. Comparing the
195 different observations, the results from A-TEEM analysis with unsupervised chemometric methods
196 were rather encouraging, but it was necessary to seek improvement through supervised classification
197 with machine learning modelling.

198 *3.2 Classification of A-TEEM data with XGBDA*

199 The multi-block A-TEEM data were analysed with XGBDA, a supervised machine learning
200 technique that has previously proven to be invaluable for wine classification (Ranaweera, Gilmore,
201 et al., 2021a, 2021b). Samples were first categorised according to the year of vintage, with Fig. 3a
202 showing the class cross-validation (CV) predictive probability according to XGBDA. This supervised
203 approach accurately classified the wines into their correct vintage year, in good accord with the k-
204 means cluster analysis (Fig. 1). In comparison, the study by Geană et al. (2019) used UV-Vis and
205 Fourier-transform infrared spectroscopy (FTIR) with LDA, obtaining overall 70% correct
206 classification for year of harvest, with young wines tending to be misclassified. The present outcome
207 with XGBDA reinforced the notion that variation among the vintages was strongly embedded in the
208 A-TEEM data, as highlighted for the k-means cluster analysis.

209 Subsequently, multi-block A-TEEM data were modelled according to subregion without
210 considering vintage year. As visualised in the class CV predictive probability from XGBDA, samples
211 from the five Barossa Valley subregions and from Eden Valley were differentiated with perfect
212 accuracy (Fig. 3b). The classification model yielded the highest possible outcome for performance
213 characteristics; i.e., sensitivity (100%), specificity (100%), error (0%), precision (100%) and F score
214 (1). This result reinforced the effectiveness of the A-TEEM technique and XGBDA algorithm for
215 successful classification of wine according to region of origin, and most impressively, at a subregional
216 level in this case. In comparison to the other spectroscopic studies on subregional classification of
217 wine, the present results stand out for their unrivalled accuracy. For example, UV, Vis and near-

infrared (NIR) spectral analyses were assessed for classification of Spanish wines belonging to four subzones within a controlled designation of origin (DO), Rías Baixas (Martelo-Vidal et al., 2013), with LDA and SVM modelling affording 85% overall correct classification for all subregions. Another study from Spain involving nuclear magnetic resonance (NMR) analysis to investigate terroir and differentiate wines of subregions in the Rioja DO achieved 92.3% correct classification using interval extended canonical variate analysis (iECVA) (López-Rituerto et al., 2012). Despite the differences in classification accuracy, it is evident from these results that an influence of terroir on wine composition can be captured from spectroscopic techniques, even for geographically close regions.

With outstanding classification results when considering either vintage or subregion, samples were further analysed considering both variables. The class CV predictive probability (Fig. 4) and the classification performance of the XGBDA model presented in Table A.3 of Appendix A show an auspicious result, with only three samples being misclassified out of 186 (SG from 2018 as WR from 2018, and NG and EV from 2020 as NG 2019 and WR 2020, respectively), giving an overall accuracy of 98.4%. The outcome was reminiscent of the previous results (Ranaweera, Gilmore, et al., 2021a, 2021b) and reiterates the power of the A-TEEM and XGBDA technique. Overall, the subregional classification was quite remarkable, achieving 100% sensitivity (true positive rate) for all subregions except for three misclassified samples (SG from 2018, 87.5%; NG from 2020, 90.09%; and EV from 2020, 90.09%). Furthermore, an error rate of 0.54% for misclassified samples was small, especially considering the low number of samples in a class (Table A.3).

The present results tended to substantiate previous studies considering regional/subregional variations among Australian Shiraz wines based on sensory or chemical analyses. Despite similarities in the sensory profiles of Shiraz wines within a GI, Johnson et al. (2013) reported that some wines still showed variability and suggested that variations in mesoclimate and geography could make it difficult to categorise those regional wines with common sensory properties. In a study by Kustos et al. (2020), sensory assessment and chemical analysis were used to investigate the regional and

subregional variations of Australian Shiraz wines, achieving 84% accuracy in cross-validation for predicting subregion with stepwise discriminant analysis using a number of chemical and sensory variables. Those authors explained that subtle variations within a region were difficult to fully identify from the sensory analysis and volatile composition because of intraregional similarities in climate and winemaking techniques. In contrast, it is clear that the sensitivity of the A-TEEM technique with XGBDA modelling has been able to identify variations among wines at a subregional level with 100% confidence based on differences in molecular fingerprint that reflect variation in chemical composition, which would be expressed in sensory profiles of the wines.

Rounding out this discussion, a recent study based on biophysical data such as climatic and topographic indices along with soil properties conducted by Bramley and Ouzman (2021) revealed that the Barossa Zone can be clustered into subregions as proposed in the Barossa Terroir project (Robinson & Sandercock, 2014). Specifically, when considering only the vineyards rather than whole area, analysis of soil properties yielded greater separation of subregions within the Barossa GI. The possibility of using the A-TEEM technique and machine learning to identify intricate variations at a subregional level should also now be considered, not only to authenticate the origin of wine at finer level, but also to potentially understand the role of subregional terroir in wine chemical composition and perceived sensory properties.

3.3 Chemical data prediction models

Other than its application for the discrimination of wine, A-TEEM has been applied to the prediction of phenolic compound concentrations in red wine using PLS regression (Ranaweera, Gilmore, et al., 2021b). This is potentially powerful due to the rapid and simple approach of the A-TEEM technique. This was expanded further in the present work through an evaluation of the simultaneous prediction of several chemical parameters, in a similar manner to that undertaken with FTIR (Friedel et al., 2013). Table 1 shows the descriptive statistics for the basic chemical parameters measured for the wines using standard (reference) methods (outlined in Section 2.3). Upon verifying that the values were in agreement with those previously reported for Australian wines including

Shiraz , the measurements were used in the development of XGB regression models for prediction of chemical parameters from A-TEEM data.

The R^2 CV and RMSECV calculated for the each parameter show a satisfactory value ($R^2 >$ 0.8, RMSECV < 0.28), demonstrating the relative robustness of the models as shown in the Fig. 5. The residual predictive deviation (RPD) values obtained for the parameters were: alcohol, 2.8; Glu+Fru, 2.9; pH, 2.1; TA, 2.4; and VA, 2.8. RPD values close to 3 indicate the reliability of the models for prediction, meaning that the present models could be applied to the quantitative determination of chemical parameters such as alcohol, glucose and fructose, and volatile acidity based on A-TEEM spectral data. The calibration results were comparable with studies using NIR spectroscopic methods for the prediction of wine chemical parameters – although the performance was slightly lower than reported by Cozzolino et al. (2011), they were slightly better than Canal and Ozen (2017). This potentially extends the utility of A-TEEM for wine analysis, but calibration of the model with many more samples and validation with an external data set to verify accuracy is required before implementing this method as a robust analytical technique that helps underpin decisions related to winemaking process or quality. Conceivably, the approach could be applied for other parameters such as tannin and organic acid concentrations upon correlation with A-TEEM data.

4. Conclusion

A combination of EEM and absorbance data from the A-TEEM technique was employed in a novel approach to classify Shiraz wine samples originating from five different subregions of Barossa Valley along with Eden Valley across three vintages, using multi-block chemometric tools. Unsupervised exploratory analysis with k-means clustering revealed the existence of clear vintage variations among the samples. Based on PCA for each separate vintage, the distribution patterns of samples according to subregion started to become evident, especially for samples from 2018. Notably, samples from the separate GI of Eden Valley were often resolved from the remainder originating from the Barossa Valley GI. Switching to supervised analysis with machine learning based on the XGBDA algorithm afforded 100% accuracy in classification according to subregion or vintage, and 98.4%

when modelling both variables, which is an impressive outcome for subregional authentication in comparison to other possible approaches. Indeed, the ability of A-TEEM molecular fingerprints to reveal subtle differences in chemical composition could assist in the verification of Australia's unique terroirs and provide novel understanding of distinctive sensory outcomes according to subregional variations in composition. Beyond authentication, the A-TEEM method was assessed as a rapid analytical technique for the simultaneous prediction of basic chemical parameters using XGB regression. R^2CV (> 0.8) and error (RMSECV < 0.28) values from the modelling were favourable, highlighting a potential alternative to the traditional laboratory methods. Thus, the versatility, simplicity and effectiveness of the method may be of benefit to the wine industry when analysing the basic parameters of numerous samples during vintage, although further validation with more samples would be required to further improve the models.

Declaration of interest

R.K.R.R., S.E.P.B., D.L.C. and D.W.J declare no conflict of interest. A.M.G., as an employee of HORIBA Instruments Inc., has no competing financial interests to declare. HORIBA Instruments Inc. did not influence the interpretation of results or decision to publish the manuscript.

Acknowledgments

We acknowledge the provision of wine samples from the Wine Australia funded Barossa Valley Terroir Project (UA1602) administered at the University of Adelaide. We are thankful to Dr. Lukas Danner, University of Adelaide for the support with sample collection. R.K.R.R. is appreciative of the financial support provided by an Australian Government Scholarship and Wine Australia supplementary scholarship (WA Ph1909). The Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au; project number IC170100008) is funded by the Australian Government with additional support from Wine Australia, Waite Research Institute, and industry partners. The University of Adelaide is a member of the Wine Innovation Cluster.

Appendix A. Supplementary data

The following is the supplementary data related to this article:

322 **References**

323 Aceto, M., Gulino, F., Calà, E., Robotti, E., Petrozziello, M., Tsolakis, C., & Cassino, C. (2020).
324 Authentication and traceability study on Barbera d'Asti and Nizza DOCG wines: The role of
325 trace- and ultra-trace elements. *6*(4), 63. https://www.mdpi.com/2306-5710/6/4/63

326 Bastian, S. E. P., & Iland OAM, P. G. (2019). Australian wine's taste evolution. In H. L. Meiselman
327 (Ed.), *Handbook of Eating and Drinking: Interdisciplinary Perspectives* (pp. 1-31). Springer
328 International Publishing. https://doi.org/10.1007/978-3-319-75388-1_169-1

329 Bekkers, T. (2012). Regional profiling: The McLaren Vale experience. *Wine and Viticulture Journal*,
330 *4*, 61–63.

331 Bramley, R. G. V., & Ouzman, J. (2021). Underpinning terroir with data: on what grounds might
332 subregionalisation of the Barossa Zone geographical indication be justified? *Australian*
333 *Journal of Grape and Wine Research*, *27*(3).
334 https://doi.org/https://doi.org/10.1111/ajgw.12513

335 Bramley, R. G. V., Ouzman, J., & Trought, M. C. T. (2020). Making sense of a sense of place:
336 precision viticulture approaches to the analysis of terroir at different scales. *OENO One*, *54*(4),
337 903-917. https://doi.org/10.20870/oeno-one.2020.54.4.3858

338 Canal, C., & Ozen, B. (2017). Monitoring of wine process and prediction of its parameters with mid-
339 infrared spectroscopy. *Journal of Food Prcoess Engineering, 40*(1), e12280.
340 https://doi.org/https://doi.org/10.1111/jfpe.12280

341 Chen, L., Capone, D. L., Nicholson, E. L., & Jeffery, D. W. (2019). Investigation of intraregional
342 variation, grape amino acids, and pre-fermentation freezing on varietal thiols and their
343 precursors for *Vitis vinifera* Sauvignon blanc. *Food Chemistry*, *295*, 637-645.
344 https://doi.org/https://doi.org/10.1016/j.foodchem.2019.05.126

345 Coetzee, P. P., van Jaarsveld, F. P., & Vanhaecke, F. (2014). Intraregional classification of wine via
346 ICP-MS elemental fingerprinting. *Food Chemistry*, *164*, 485-492.
347 https://doi.org/https://doi.org/10.1016/j.foodchem.2014.05.027

348 Cozzolino, D., Cynkar, W., Shah, N., & Smith, P. (2011). Feasibility study on the use of attenuated
349 total reflectance mid-infrared for analysis of compositional parameters in wine. *Food*
350 *Research International*, *44*(1), 181-186.
351 https://doi.org/https://doi.org/10.1016/j.foodres.2010.10.043
352 Cozzolino, D., Cynkar, W. U., Shah, N., Dambergs, R. G., & Smith, P. A. (2009). A brief introduction
353 to multivariate methods in grape and wine analysis. *International Journal of Wine Research*,
354 123. https://doi.org/10.2147/ijwr.s4585
355 Friedel, M., Patz, C.-D., & Dietrich, H. (2013). Comparison of different measurement techniques and
356 variable selection methods for FT-MIR in wine analysis. *Food Chemistry*, *141*(4), 4200-4207.
357 https://doi.org/https://doi.org/10.1016/j.foodchem.2013.06.120
358 Geană, E.-I., Ciucure, C. T., Apetrei, C., & Artem, V. (2019). Application of spectroscopic UV-vis
359 and FT-IR screening techniques coupled with multivariate statistical analysis for red wine
360 authentication: Varietal and vintage year discrimination. *Molecules*, *24*(22), 4166.
361 https://www.mdpi.com/1420-3049/24/22/4166
362 Gilmore, A. (2014). How to collect national institute of standards and technology (NIST) traceable
363 fluorescence excitation and emission spectra. In Y. Engelborghs & A. J. W. G. Visser (Eds.),
364 *Fluorescence Spectroscopy and Microscopy: Methods and Protocols* (pp. 3-27). Humana
365 Press. https://doi.org/10.1007/978-1-62703-649-8_1
366 Gilmore, A., Akaji, S., & Csatorday, K. (2017). Spectroscopic analysis of red wines with A-TEEM
367 molecular fingerprinting. *Readout*, *2*(49), 41-48.
368 https://www.horiba.com/in/publications/readout/article/spectroscopic-analysis-of-red-wines-
369 with-a-teem-molecular-fingerprinting-53035/
370 Johnson, T. E., Hasted, A., Ristic, R., & Bastian, S. E. P. (2013). Multidimensional scaling (MDS),
371 cluster and descriptive analyses provide preliminary insights into Australian Shiraz wine
372 regional characteristics. *Food Quality and Preference*, *29*(2), 174-185.
373 https://doi.org/https://doi.org/10.1016/j.foodqual.2013.03.010

374 Kustos, M., Gambetta, J. M., Jeffery, D. W., Heymann, H., Goodman, S., & Bastian, S. E. P. (2020).
375 A matter of place: Sensory and chemical characterisation of fine Australian Chardonnay and
376 Shiraz wines of provenance. *Food Research International*, *130*, 108903.
377 https://doi.org/https://doi.org/10.1016/j.foodres.2019.108903
378 López-Rituerto, E., Savorani, F., Avenoza, A., Busto, J. H., Peregrina, J. M., & Engelsen, S. B.
379 (2012). Investigations of La Rioja terroir for wine production using ^{1}H NMR metabolomics.
380 *Journal of Agricultural and Food Chemistry*, *60*(13), 3452-3461.
381 https://doi.org/10.1021/jf204361d
382 Lorrain, B., Chira, K., & Teissedre, P.-L. (2011). Phenolic composition of Merlot and Cabernet-
383 Sauvignon grapes from Bordeaux vineyard for the 2009-vintage: Comparison to 2006, 2007
384 and 2008 vintages. *Food Chemistry*, *126*(4), 1991-1999.
385 https://doi.org/https://doi.org/10.1016/j.foodchem.2010.12.062
386 Martelo-Vidal, M. J., Domínguez-Agis, F., & Vázquez, M. (2013). Ultraviolet/visible/near-infrared
387 spectral analysis and chemometric tools for the discrimination of wines between subzones
388 inside a controlled designation of origin: a case study of Rías Baixas. *Australian Journal of*
389 *Grape and Wine Research*, *19*(1), 62-67. https://doi.org/https://doi.org/10.1111/ajgw.12003
390 Martin, A. E., Watling, R. J., & Lee, G. S. (2012). The multi-element determination and regional
391 discrimination of Australian wines. *Food Chemistry*, *133*(3), 1081-1089.
392 http://www.sciencedirect.com/science/article/pii/S0308814612001781
393 Niimi, J., Liland, K. H., Tomic, O., Jeffery, D. W., Bastian, S. E. P., & Boss, P. K. (2021). Prediction
394 of wine sensory properties using mid-infrared spectra of Cabernet Sauvignon and Chardonnay
395 grape berries and wines. *Food Chemistry*, *344*, 128634.
396 https://doi.org/https://doi.org/10.1016/j.foodchem.2020.128634
397 Niimi, J., Tomic, O., Næs, T., Bastian, S. E. P., Jeffery, D. W., Nicholson, E. L., Maffei, S. M., &
398 Boss, P. K. (2020). Objective measures of grape quality: From Cabernet Sauvignon grape

399 composition to wine sensory characteristics. *LWT-Food Science and Technology*, *123*,
400 109105. https://doi.org/https://doi.org/10.1016/j.lwt.2020.109105
401 Pearson, W., Schmidtke, L. M., Francis, I. L., Li, S., Hall, A., & Blackman, J. W. (2021). Regionality
402 in Australian Shiraz: Compositional and climate measures that relate to key sensory attributes.
403 *Australian Journal of Grape and Wine Research*, *27*(3).
404 https://doi.org/https://doi.org/10.1111/ajgw.12499
405 Ranaweera, R. K. R., Capone, D. L., Bastian, S. E. P., Cozzolino, D., & Jeffery, D. W. (2021). A
406 review of wine authentication using spectroscopic approaches in combination with
407 chemometrics. *Molecules*, *26*(14), 4334. https://www.mdpi.com/1420-3049/26/14/4334
408 Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021a).
409 Authentication of the geographical origin of Australian Cabernet Sauvignon wines using
410 spectrofluorometric and multi-element analyses with multivariate statistical modelling. *Food*
411 *Chemistry*, *335*, 127592. https://doi.org/https://doi.org/10.1016/j.foodchem.2020.127592
412 Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021b).
413 Spectrofluorometric analysis combined with machine learning for geographical and varietal
414 authentication, and prediction of phenolic compound concentrations in red wine. *Food*
415 *Chemistry*, *361*, 130149. https://doi.org/https://doi.org/10.1016/j.foodchem.2021.130149
416 Riovanto, R., Cynkar, W. U., Berzaghi, P., & Cozzolino, D. (2011). Discrimination between Shiraz
417 wines from different Australian regions: The role of spectroscopy and chemometrics. *Journal*
418 *of Agricultural and Food Chemistry*, *59*(18), 10356-10360. https://doi.org/10.1021/jf202578f
419 Robinson, S., & Sandercock, N. (2014). *An analysis of climate, soil and topographic information to*
420 *aid the understanding of Barossa terroir*. PIRSA Spatial Information Services, Government
421 of South Australia, Adelaide. https://www.barossawine.com/vineyards/barossa-
422 grounds/barossa-grounds-toolkit/
423 Roullier-Gall, C., Boutegrabet, L., Gougeon, R. D., & Schmitt-Kopplin, P. (2014). A grape and wine
424 chemodiversity comparison of different appellations in Burgundy: Vintage vs terroir effects.

425 *Food Chemistry*, *152*, 100-107.
426 https://doi.org/https://doi.org/10.1016/j.foodchem.2013.11.056
427 Sáenz-Navajas, M.-P., & Jeffery, D. W. (2021). Perspectives on wines of provenance: Sensory
428 typicality, quality, and authenticity. *ACS Food Science & Technology*, *1*(6), 986-992.
429 https://doi.org/10.1021/acsfoodscitech.1c00128
430 Souza Gonzaga, L., Bastian, S. E. P., Capone, D. L., Ranaweera, R. K. R., & Jeffery, D. W. (2021).
431 Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *OENO*
432 *One*, *55*(4), 19-33. https://doi.org/10.20870/oeno-one.2021.55.4.4805
433 Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Defining wine typicity:
434 sensory characterisation and consumer perspectives. *27*(2), 246-256.
435 https://doi.org/https://doi.org/10.1111/ajgw.12474
436 van Leeuwen, C., Friant, P., Choné, X., Tregoat, O., Koundouras, S., & Dubourdieu, D. (2004).
437 Influence of climate, soil, and cultivar on terroir. *American Journal of Enology and*
438 *Viticulture*, *55*(3), 207-217. https://www.scopus.com/inward/record.uri?eid=2-s2.0-
439 5344244786&partnerID=40&md5=f62e072f81e9b9cf5435f4ef143a627f
440 Werner, M., & Roche, M. (2016). *Clare Valley rocks - The earth beneath our vines*. Retrieved
441 20/06/2021 from https://doi.org/10.13140/RG.2.2.35230.54086.
442 Wine Australia. (2021). *Geographical indications*. Retrieved 10.06.2021 from
443 https://www.wineaustralia.com/labelling/register-of-protected-gis-and-other-
444 terms/geographical-indications
445 Xie, H., Konate, M., Sai, N., Tesfamicael, K. G., Cavagnaro, T., Gilliham, M., Breen, J., Metcalfe,
446 A., Stephen, J. R., De Bei, R., Collins, C., & Lopez, C. M. R. (2017). Global DNA methylation
447 matterns can play a role in defining terroir in grapevine *(Vitis vinifera cv.* Shiraz). *8*(1860).
448 https://doi.org/10.3389/fpls.2017.01860
449 Zhou, J., Cavagnaro, T. R., De Bei, R., Nelson, T. M., Stephen, J. R., Metcalfe, A., Gilliham, M.,
450 Breen, J., Collins, C., & López, C. M. R. (2021). Wine terroir and the soil bacteria: An

451 amplicon sequencing–based assessment of the Barossa Valley and its sub-regions. *11*(3358).

452 https://doi.org/10.3389/fmicb.2020.597944

Figure Captions

Fig. 1. K-means partitional cluster analysis (k = 3) of A-TEEM data for subregional Shiraz wines (n = 186) from three different vintages (Class 1: 2018 in green; Class 2: 2019 in red; Class 3: 2020 in blue).

Fig. 2. Scores for the wine samples from PCA of multi-block data according to subregion for individual vintage years, showing **a**) 2018, **b**) 2019, and **c**) 2020. CG, Central Grounds (blue); EE, Eastern Edge (purple); EV, Eden Valley (green); NG, Northern Grounds (pink); SG, Southern Grounds (yellow); WR, Western Ridge (brown).

Fig. 3. Class CV predicted with XGBDA classification using A-TEEM data for subregional Shiraz wines (n = 186) according to **a**) vintage year for wine produced in 2018 (orange triangles), 2019 (purple triangles) and 2020 (green stars), and **b**) subregion within the Barossa Valley GI along with Eden Valley GI for all three vintages. CG, Central Grounds (blue triangles); EE, Eastern Edge (purple triangles); EV, Eden Valley (green stars); NG, Northern Grounds (pink circles); SG, Southern Grounds (yellow diamonds); WR, Western Ridge (brown squares).

Fig. 4. Class CV predicted by region/subregion across three vintages (2018–2020) from XGBDA modelling of A-TEEM data for Shiraz wines (n = 186) from different subregions within the Barossa GI and from Eden Valley GI, South Australia. CG, Central Grounds; EE, Eastern Edge; EV, Eden Valley; NG, Nothern Grounds; SG, Southern Grounds; WR, Western Ridge.

Fig. 5. XGB regression of measured vs predicted parameters using reference methods and A-TEEM data for subregional Shiraz wines (n = 186) showing **a**) alcohol (% v/v), **b**) glucose + fructose (g L^{-1}), **c**) pH, **d**) titratable acidity (g L^{-1}), and **e**) volatile acidity (g L^{-1}) along with calibration and validation statistics.

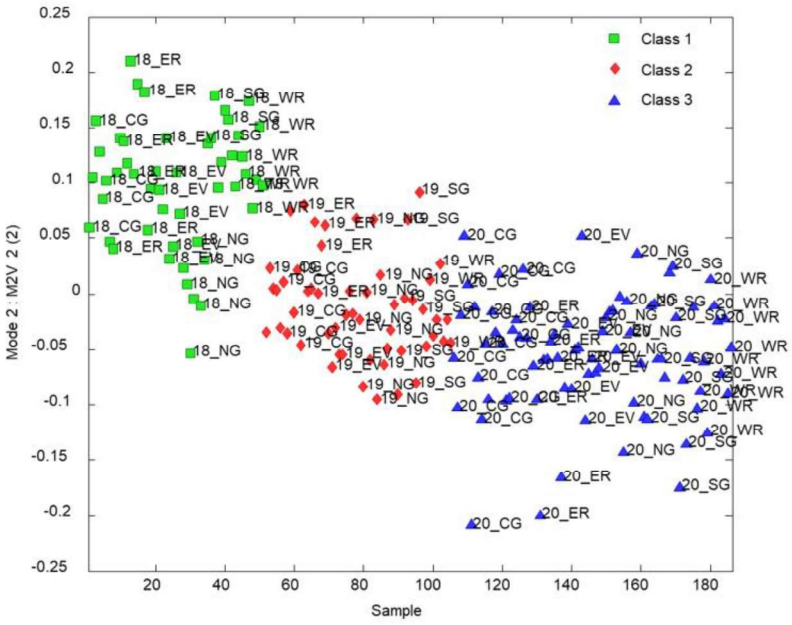

Class 1
Class 2
Class 3
Mode 2 : M2V 2 (2)
Sample
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
-0.25
20
40
60
80
100
120
140
160
180

Fig. 1

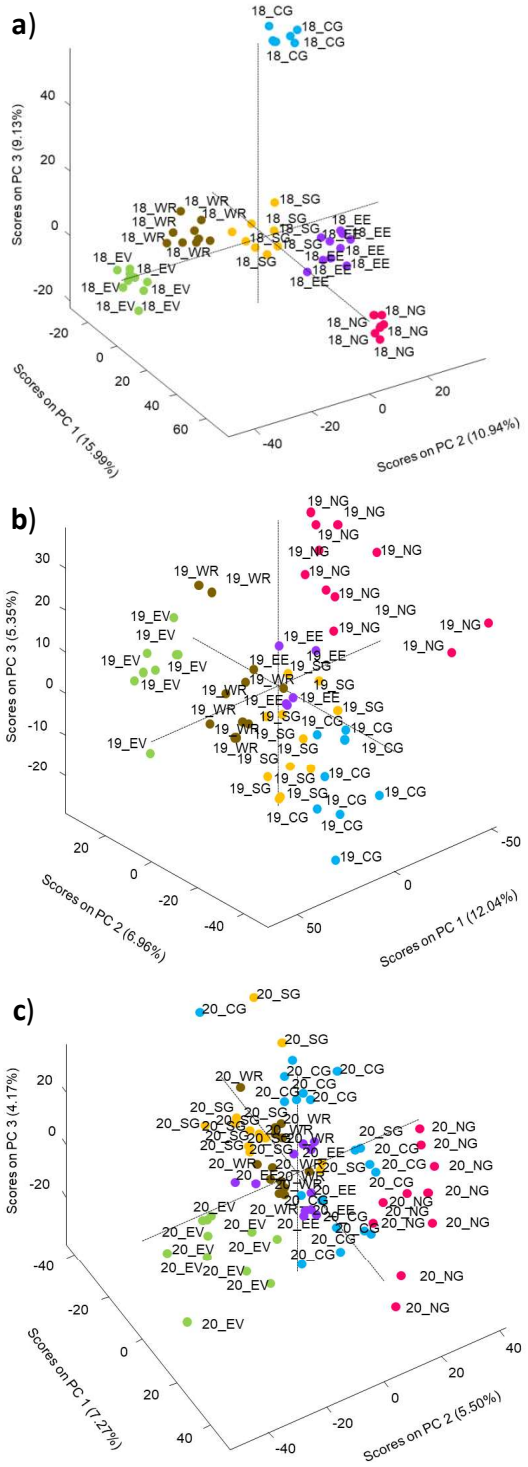

a)
Scores on PC 3 (9.13%)
Scores on PC 1 (15.99%)
Scores on PC 2 (10.94%)
b)
Scores on PC 3 (5.35%)
Scores on PC 2 (6.96%)
Scores on PC 1 (12.04%)
c)
Scores on PC 3 (4.17%)
Scores on PC 1 (7.27%)
Scores on PC 2 (5.50%)

Fig. 2

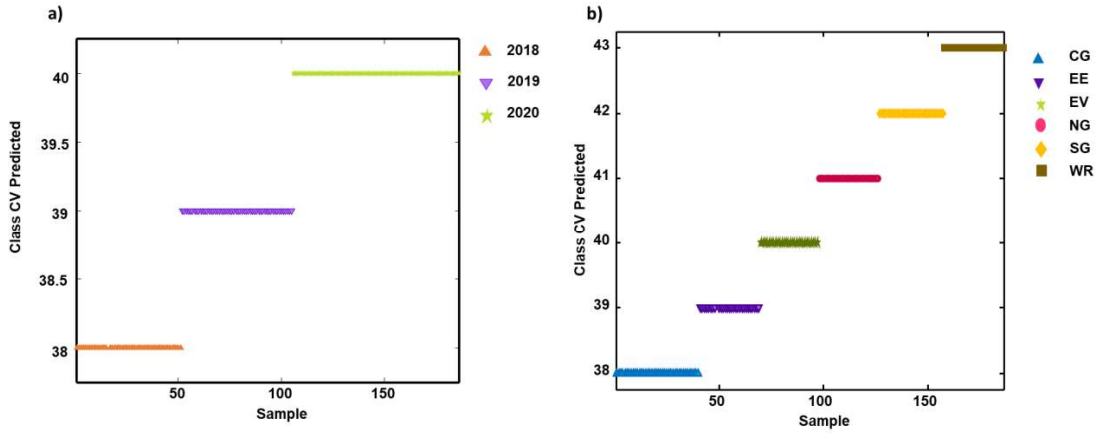

a)
Class CV Predicted
Sample
2018
2019
2020
b)
Class CV Predicted
Sample
CG
EE
EV
NG
SG
WR

Fig. 3

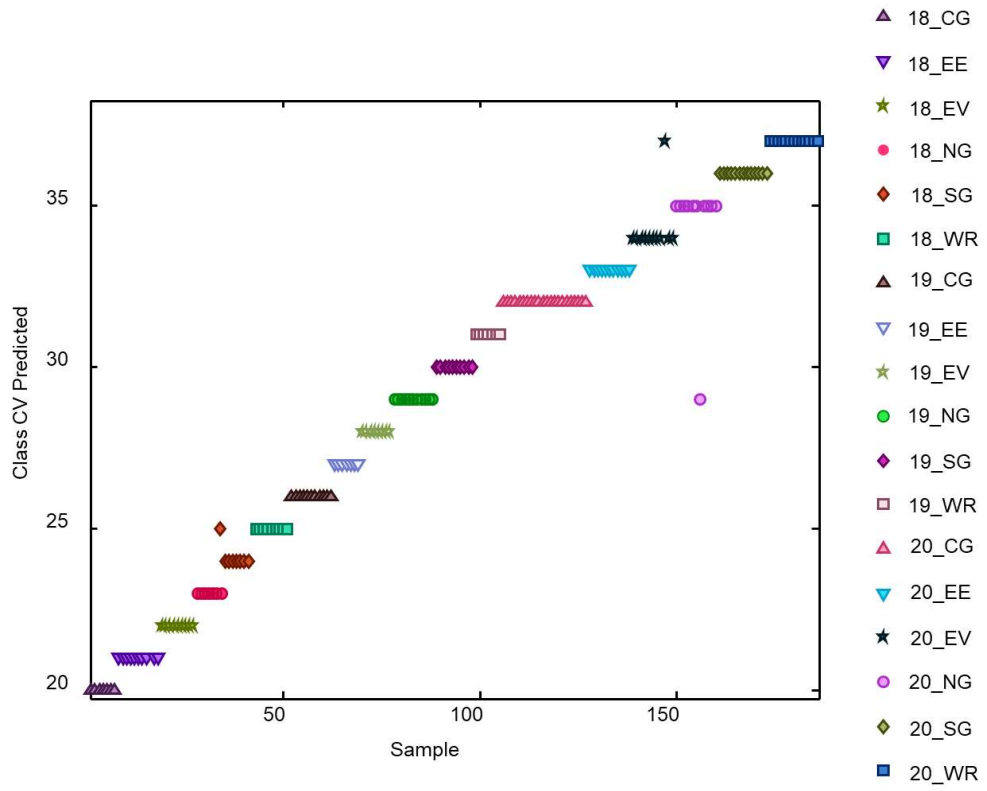

Class CV Predicted
35
30
25
20
50
100
150
Sample
18_CG
18_EE
18_EV
18_NG
18_SG
18_WR
19_CG
19_EE
19_EV
19_NG
19_SG
19_WR
20_CG
20_EE
20_EV
20_NG
20_SG
20_WR

Fig. 4

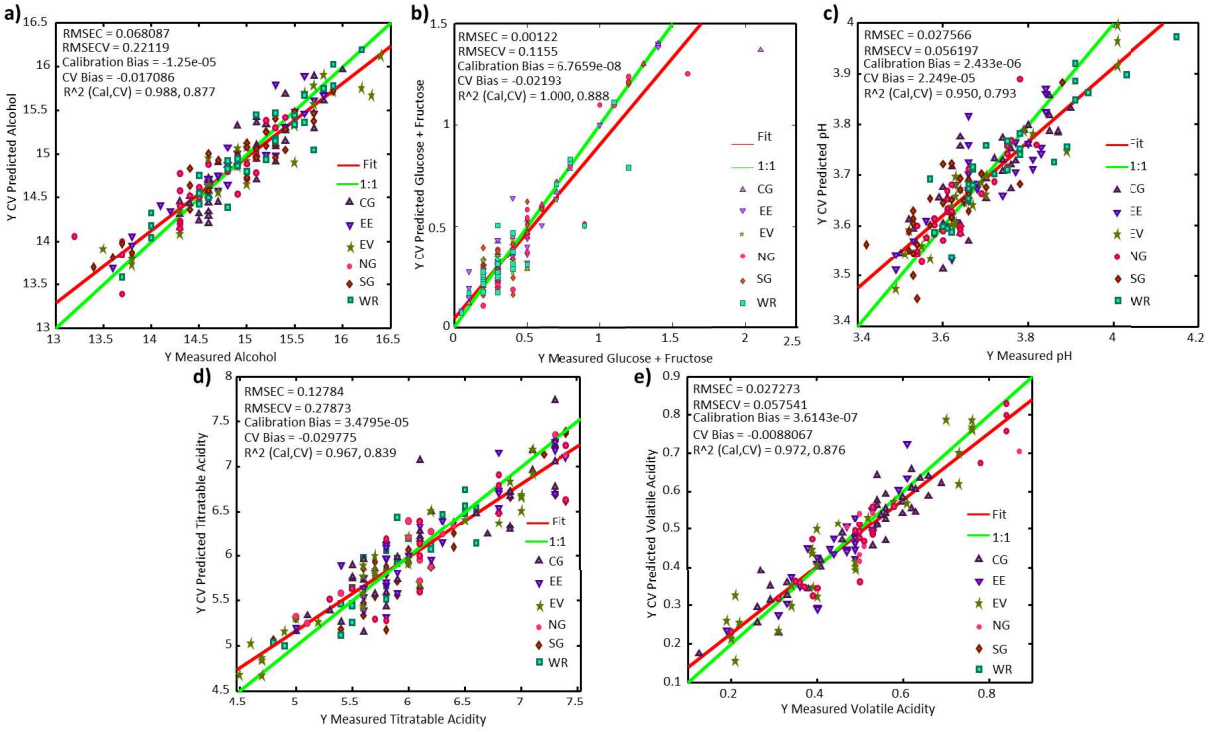

a)
RMSEC = 0.068087
RMSECV = 0.22119
Calibration Bias = -1.25e-05
CV Bias = -0.017086
R^2 (Cal,CV) = 0.988, 0.877
Y CV Predicted Alcohol
Y Measured Alcohol
b)
RMSEC = 0.00122
RMSECV = 0.1155
Calibration Bias = 6.7659e-08
CV Bias = -0.02193
R^2 (Cal,CV) = 1.000, 0.888
Y CV Predicted Glucose + Fructose
Y Measured Glucose + Fructose
c)
RMSEC = 0.027566
RMSECV = 0.056197
Calibration Bias = 2.433e-06
CV Bias = 2.249e-05
R^2 (Cal,CV) = 0.950, 0.793
Y CV Predicted pH
Y Measured pH
d)
RMSEC = 0.12784
RMSECV = 0.27873
Calibration Bias = 3.4795e-05
CV Bias = -0.029775
R^2 (Cal,CV) = 0.967, 0.839
Y CV Predicted Titratable Acidity
Y Measured Titratable Acidity
e)
RMSEC = 0.027273
RMSECV = 0.057541
Calibration Bias = 3.6143e-07
CV Bias = -0.0088067
R^2 (Cal,CV) = 0.972, 0.876
Y CV Predicted Volatile Acidity
Y Measured Volatile Acidity
Fit
1:1
CG
EE
EV
NG
SG
WR

Fig. 5

Table 1

Descriptive statistics for basic chemical parameters of Shiraz wines (n = 186) from Barossa subregions for vintages 2018, 2019, and 2020.

Statistic	Alcohol (% v/v)	Glu + Fruc (g L^{-1})	pH	Titratable acidity (g L^{-1})	Volatile acidity (g L^{-1})
Minimum	13.2	0.05	3.42	4.5	0.13
Maximum	16.4	2.10	4.15	7.4	0.87
Range	3.2	2.05	0.73	2.9	0.75
1st Quartile	14.6	0.23	3.60	5.6	0.38
Median	14.9	0.30	3.66	6.0	0.50
3rd Quartile	15.3	0.50	3.75	6.6	0.61
Mean	14.9	0.458	3.68	6.1	0.50
Variance (n−1)	0.39	0.117	0.02	0.48	0.03
Standard deviation (n−1)	0.63	0.342	0.12	0.69	0.16
Variation coefficient (n−1)	0.04	0.747	0.03	0.11	0.33

APPENDIX A

SUPPLEMENTARY DATA FOR

Absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine

Ranaweera K. R. Ranaweera[a], Susan E. P. Bastian[a,b], Adam M. Gilmore[c], Dimitra L. Capone[a,b], David W. Jeffery[a,b,*]

[a] Department of Wine Science and Waite Research Institute, The University of Adelaide (UA), PMB 1, Glen Osmond, South Australia 5064, Australia

[b] Australian Research Council Training Centre for Innovative Wine Production, UA, PMB 1, Glen Osmond, South Australia 5064, Australia

[c] HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, USA

[*] Corresponding author.

E-mail address: david.jeffery@adelaide.edu.au (D.W. Jeffery).

Table of Contents

Page

Table A.1. Summary of the different vintages and region/subregions of the Shiraz wines. S-2

Table A.2. Climatic data of the Barossa regions for year 2017-2020. S-2

Table A.3. Confusion matrix results for XGBDA analysis of multi-block data (EEM and absorbance) for Shiraz wines modelled according to subregion and vintage. S-2

Table A.4. Least squares means and one-way ANOVA results of basic chemical measures for the Shiraz wines. S-3

Fig. A.1. Locations of the subregions for the samples collected from the Barossa GI, South Australia. S-3

Table A.1. Summary of the different vintages and region/subregion of the Shiraz wines from the Barossa GI analysed in this study.

	No. of samples from each vintage[b]			
Subregion	**2018**	**2019**	**2020**	**Total**
Central Grounds (CG)	7	11	22	40
Eastern Edge (EE)	9	9	11	29
Eden Valley (EV)[a]	9	8	11	28
Northern Grounds (NG)	7	11	11	29
Southern Grounds (SG)	8	10	13	31
Western Ridge (WR)	9	7	13	29
Total	49	56	81	186

[a]Eden Valley is a region according to the Australian GI system. [b]Unreleased commercial wines included: 2 SG and 3 EE in 2018; 3 SG, 5 CG, and 2 EE in 2019; 1 EV, 3 SG, 1 NG, 5 CG, and 1 WR in 2020.

Table A.2. Climatic data of the Barossa and Eden Valley regions for years 2017-2020, sourced from Wine Australia Marketing Insights (https://www.wineaustralia.com/market-insights/regional-snapshots).

	Mean January Temperature (℃)		**Annual Rainfall (mm)**		**Growing Degree Days (GDD)**	
Year	**Barossa Valley**	**Eden Valley**	**Barossa Valley**	**Eden Valley**	**Barossa Valley**	**Eden Valley**
2017-2018	23.7	22.8	466	495	2089	1918
2018-2019	24.8	24.4	366	347	2120	2031
2019-2020	21.0	21.1	420	386	1862	1744

Table A.3. Confusion matrix results for XGBDA analysis of multi-block data (EEM and absorbance) for Shiraz wines modelled according to region/subregion and vintage.

Class[a]	**No.**	**Sensitivity%**	**Specificity%**	**Err%**	**Precision%**	**F1**
18_CG	7	100.00	100.00	0.00	100.00	1.00
18_EE	9	100.00	100.00	0.00	100.00	1.00
18_EV	9	100.00	100.00	0.00	100.00	1.00
18_NG	7	100.00	100.00	0.00	100.00	1.00
18_SG	8	87.50	100.00	0.54	100.00	0.93
18_WR	9	100.00	99.43	0.54	90.00	0.95
19_CG	11	100.00	100.00	0.00	100.00	1.00
19_EE	9	100.00	100.00	0.00	100.00	1.00
19_EV	8	100.00	100.00	0.00	100.00	1.00
19_NG	11	100.00	99.43	0.54	91.67	0.96
19_SG	10	100.00	100.00	0.00	100.00	1.00
19_WR	7	100.00	100.00	0.00	100.00	1.00
20_CG	22	100.00	100.00	0.00	100.00	1.00
20_EE	11	100.00	100.00	0.00	100.00	1.00
20_EV	11	90.09	100.00	0.54	100.00	0.95
20_NG	11	90.09	100.00	0.54	100.00	0.95
20_SG	13	100.00	100.00	0.00	100.00	1.0
20_WR	13	100.00	99.42	0.54	92.8	0.96

[a]Region abbreviation as per Table A.1 along with vintage year corresponding to 2018, 2019, and 2020.

Table A.4. Least squares means and one-way ANOVA results (significant p-value in bold) of basic chemical measures for the region/subregion. Different letters within a column indicate statistically significant differences among the means for region/subregion (Tukey's HSD, $\alpha = 0.05$).

Subregion[a]	Alcohol (% v/v)	Glu + Fruc (g L^{-1})	pH	Titratable Acidity (g L^{-1})	Volatile Acidity (g L^{-1})
CG	14.9	0.50	3.69 ab	6.3	0.49
EE	14.8	0.50	3.70 ab	6.2	0.44
EV	15.0	0.32	3.70 ab	5.8	0.45
NG	14.7	0.56	3.64 b	6.2	0.52
SG	14.9	0.46	3.61 b	6.1	0.52
WR	15.0	0.412	3.75 a	5.9	0.53
p-value	0.277	0.147	**0.000**	0.115	0.166

[a]Abbreviations as per Table A.1.

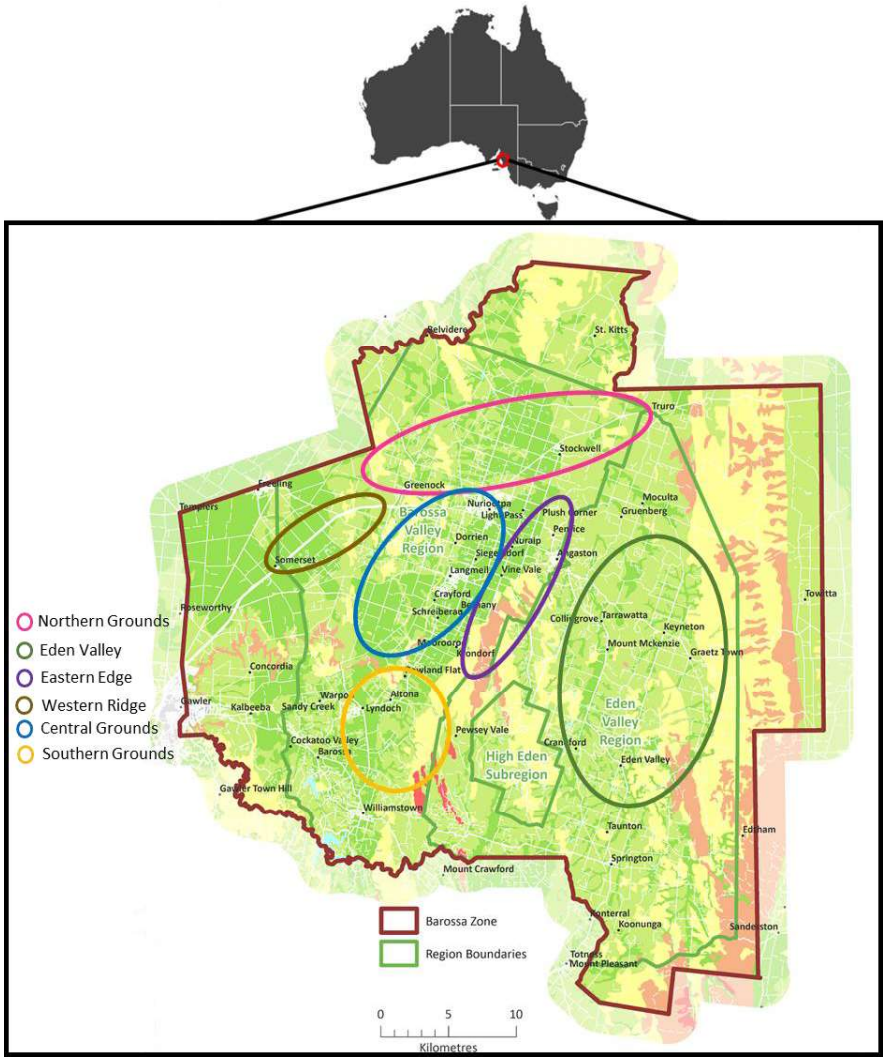
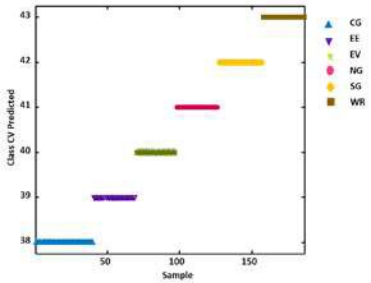


Fig. A.1. Locations of the region/subregion of the samples collected from the Barossa GI, South Australia (Map based on Robinson, S., & Sandercock, N. (2014). An analysis of climate, soil and topographic information to aid the understanding of Barossa terroir. PIRSA Spatial Information Services, Government of South Australia, Adelaide. https://www.barossawine.com/vineyards/barossa-grounds/barossa-grounds-toolkit/).

Graphical Abstract

Subregions of
Barossa Valley

Spectrofluorometric
wine analysis



XGBDA model for
subregional authentication

Declaration of Interest Statement

Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

R.K.R.R., S.E.P.B., D.L.C. and D.W.J have no interests to declare. A.M.G., as an employee of HORIBA Instruments Inc., has no competing financial interests to declare. HORIBA Instruments Inc. did not influence the interpretation of results or decision to publish the manuscript.

CRediT authorship contribution statement

R.K.R.R.: Conceptualisation, Formal analysis, Investigation, Methodology, Visualisation, Writing - original draft, Review & editing. **S.E.P.B:** Conceptualisation, Funding acquisition, Resources, Supervision, Writing – review & editing. **A.M.G:** Methodology, Formal analysis, Writing – review & editing. **D.L.C.:** Conceptualisation, Supervision, Writing – review & editing. **D.W.J.:** Conceptualisation, Funding acquisition, Project administration, Resources, Supervision, Writing - review & editing.

Chapter 6

Manuscript for Publication

Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines

Ranaweera, R. K. R., Gilmore, A. M., Bastian, S. E. P., Capone, D. L., & Jeffery, D. W. (2021). Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines. OENO One (Submitted for publication).

Statement of Authorship

Title of Paper	Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Ranaweera, R. K. R., Gilmore, A. M., Bastian, S. E. P., Capone, D. L., & Jeffery, D. W. (2021). Feasibility of using spectrofluorometric analysis to the molecular fingerprint of wine through the winemaking process and recognise the blending percentage of different varietal wines, OENO One

Principal Author

Name of Principal Author (Candidate)	Ranaweera K. R. Ranaweera		
Contribution to the Paper	Contributed to the experimental design and A-TEEM analysis for the wine samples. Carried out data collection, processing, and applied statistical techniques, created models and interpreted data for the manuscript. Drafted the manuscript, and reviewed and edited the final draft.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	04/11/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Adam M. Gilmore		
Contribution to the Paper	Contributed to design of methodology for wine analysis. Applied statistical techniques in formal analysis of the study data. Reviewed, edited and revised the manuscript for publication.		
Signature		Date	2021/11/15

Name of Co-Author	Susan E.P. Bastian		
Contribution to the Paper	Contributed to the research idea. Coordinated the wine sample acquisition. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript.		
Signature		Date	15/11/2021

Name of Co-Author	Dimitra L. Capone		
Contribution to the Paper	Contributed to the research idea. Supervised the project and oversaw research activity planning and execution. Reviewed, edited and revised the manuscript.		
Signature		Date	15/11/2021

Name of Co-Author	David W. Jeffery		
Contribution to the Paper	Original conceptualisation and design of the study. Acquired funding and provided resources for the project leading to this publication. Supervised the project and undertook project administration throughout the study. Critically evaluated, reviewed, edited and revised the manuscript.		
Signature		Date	15/11/2021

Feasibility of using spectrofluorometric analysis to trace the molecular fingerprint of wine through the winemaking process and recognise the blending percentage of different varietal wines

Ranaweera K.R. Ranaweera[1], Adam M. Gilmore[2], Susan E.P. Bastian[1,3], Dimitra L. Capone[1,3], and David W. Jeffery [1,3*]

[1] Department of Wine Science and Waite Research Institute, The University of Adelaide, PMB 1, Glen Osmond, South Australia 5064, Australia

[2] HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, USA

[3] Australian Research Council Training Centre for Innovative Wine Production, The University of Adelaide, PMB 1, Glen Osmond, South Australia 5064, Australia

*corresponding author: david.jeffery@adelaide.edu.au

ABSTRACT

As a robust analytical method, the application of spectrofluorometric analysis with machine learning modelling has recently been used to authenticate wine from different regions, vintages and varieties. Whether the molecular fingerprint obtained with this approach is maintained throughout the winemaking process has been preliminarily investigated along with an assessment of different percentages of wine in a blend. Monovarietal wine samples were collected at different stages of the winemaking process and analysed with an absorbance-transmission and fluorescence excitation emission matrix (A-TEEM) technique. Wines were clustered tightly according to origin for the different winemaking stages, with some clear separation of different regions and varieties based on principal component analysis. On the other hand, wines were classified with 100 % accuracy according to varietal origin using extreme gradient boosting (XGB) discriminant analysis. The sensitivity of the technique was such that it allowed for accurate modelling of wine blends containing as little as 1 % of Cabernet Sauvignon or Grenache in Shiraz wine when employing XGB regression, which performed

better than partial least squares regression. The overall results indicated the potential for applying A-TEEM and machine learning modelling to wine chemical traceability through production to guarantee the provenance of wine or identify the composition of a blend.

KEYWORDS

Authenticity, excitation-emission matrix, traceability, chemometrics, vinification

INTRODUCTION

Wine is an attractive target for fraud due to it being a luxury product in a high value industry, worth hundreds of billion dollars globally. Wine fraud can occur in different forms, such as dilution, substitution, illegal addition, and mislabelling (Ranaweera, Souza Gonzaga, et al., 2021). To ensure the provenance of wine and to combat wine fraud, it is important to verify the origin and identity of the product by applying proper authentication and traceability techniques. Even though a number of analytical methods have been developed for wine authentication, it is challenging to find a technique to verify the original fingerprint of the product that has been maintained throughout production, due to the complexity of the winemaking process (Aceto et al., 2013).

At the very least, winemaking involves alcoholic fermentation, but can encompass other processes such as malolactic fermentation, use of permitted additives or maturation techniques, and blending of different varietals. Each of these processes imparts alterations to wine composition: alcoholic fermentation produces compounds such as higher alcohols, esters, glycerol, acetaldehyde, and acids (Styger et al., 2011); malolactic fermentation involves changes that impact wine aroma and flavour profiles beyond the conversion of malic acid into lactic acid (Lonvaud-Funel, 2010); and interactions of wine macromolecules such as polysaccharides with aroma compounds, tannins, and proteins also affect the wine matrix (Jones-Moore et al., 2022). Some components in wine do not change significantly during the vinification process, however, which offers the opportunity to identify chemical markers that could be applied for authentication purpose (Catalano et al., 2016; Versari et al., 2014).

Few studies have been conducted to verify the possibility of tracing chemical markers during winemaking. Analysis of metal composition throughout the winemaking process has identified only a few elements that maintained constant concentrations (Castiñeira et al., 2004). In their study, Almeida and Vasconcelos (2004) have shown that $^{87}Sr/^{86}Sr$ isotope values were statistically identical and can be applied to the provenance of soil, and respective grape juice and wine. A study of wine phenolic profile during winemaking using Fourier-transform infrared spectroscopy has identified that total phenolic content did not change significantly after primary

and malolactic fermentations (Preserova et al., 2015). However, studies have shown that the blending process used to produce a finished wine can affect polyphenols and colour (Li et al., 2020) and bentonite used for protein stabilisation can influence the distribution of various metals (Aceto et al., 2013). Furthermore, although blending is an important step for producing wine with appealing sensory properties (Dooley et al., 2012) that may underpin the reputation of a designated origin (DO), such as Bordeaux blends involving Cabernet and Merlot or Australian Shiraz and Cabernet blends (Souza Gonzaga et al., 2021; Wine Australia, 2017), it can introduce uncertainty for confirming authenticity. For example, there could be unauthorised blending of DO wine with a small percentage of non-DO wine to increase total volume, or there may be a need to identify blending proportions for labelling requirements, such as having 85 % or more of the variety or geographic indication stated on the bottle label in accordance with the label integrity programme in Australia (Wine Australia, 2018). Imparato et al. (2011) applied nuclear magnetic resonance (NMR) profiling to a range of red wine varieties and achieved a precision of about 10 % when differentiating wine blends. For authentication purposes, however, a robust (and preferably rapid) method with high accuracy was still required, to verify the blends of different grape varieties.

Considering that fluorescence spectroscopy can offer a viable method for wine authentication (Ranaweera, Gilmore, et al., 2021a, 2021b), the present study used a spectrofluorometric technique (absorbance-transmission and fluorescence excitation emission matrix, or A-TEEM) in combination with machine learning modelling to test two hypotheses for the first time: 1) the molecular fingerprint of wine as a function of origin can be traced through steps of the winemaking process, and 2) the blending percentages of different wines can be detected. The effectiveness of the cross-validated models was evaluated and compared according to the score probabilities in the confusion matrix and root mean square error of cross-validation (RMSECV) along with coefficient of determination of cross-validation (R^2 CV).

MATERIALS AND METHODS

1. Chemicals and solvents

HPLC gradient grade absolute ethanol and analytical grade 37 % hydrochloric acid (HCl) were purchased from Chem-Supply (Port Adelaide, SA, Australia). High purity water was obtained from a Milli-Q purification system (Millipore, North Ryde, NSW, Australia).

2. Wine samples

Two sets of wine samples were obtained to examine the stage of wine production and for blending experiments. For stage of production, five different monovarietal wines (Grenache

from Alverstoke vineyard and from Coombe vineyard at the University of Adelaide's Waite Campus, Mataro from Coombe vineyard, Shiraz from Barossa Valley, and Nebbiolo from Southern Flinders Ranges) were collected in 2021 from the research and teaching winery at the Waite Campus at three different processing stages: post primary fermentation (PF) when glucose and fructose were less than 2 g/L; post malolactic fermentation (MF) when malic acid concentration was less than 0.1 g/L; and pre-blending (PB) from 225 L barrels. For blending experiments, three different commercially produced but unreleased monovarietal wines (Shiraz from Langhorne Creek, Cabernet Sauvignon from Langhorne Creek, and Grenache from Riverland) were obtained from a local producer in 2020.

3. Analytical procedures for basic chemical parameters

Wine pH and titratable acidity (TA) were measured with an autotitrator and alcohol content (percentage by volume) was measured with a density meter by Commercial Services at the Australian Wine Research Institute. Analyses were undertaken in duplicate.

4. Sample preparation and A-TEEM analysis for winemaking stages

Samples were obtained from fermentation vessels or barrels at PF, MF, and PB stages of production and stored in plastic containers in a freezer at −20 °C until required for analysis to inhibit fermentation. At the time of analysis, samples were defrosted at room temperature and prepared and analysed in duplicate as described by Ranaweera, Gilmore, et al. (2021b), undertaking two measurements of each replicate sample. Briefly, samples (1 mL) were centrifuged (Eppendorf 5415D, Adelab Scientific, Thebarton, SA, Australia) at 9300 × g for 10 min and an aliquot (40 µL) was diluted 1:100 with 50 % aqueous ethanol that had been adjusted to pH 2 with HCl and degassed by vacuum filtration (0.45 µm PTFE membrane). The dilution factor of wine-to-solvent was determined by considering the absorbance values of samples according to Beer-Lambert law (Gilmore, 2014). Samples were mixed for 60 s using a benchtop vortex (Grant-bio, PV-1) and degassed by sonication for 10 min with a Unisonics ultrasonic cleaner (Rowe Scientific, Adelaide, SA, Australia). A-TEEM analysis was conducted with a HORIBA Scientific Aqualog spectrophotometer (version 4.2, Quark Photonics, Adelaide, SA, Australia) using the same instrument settings as reported previously (Ranaweera, Gilmore, et al., 2021b) (i.e., excitation wavelength range of 240–800 nm with a 5 nm increment under medium gain and 0.2 s integration time; emission wavelength range of 242–824 nm with a 4.66 nm increment as set by the instrument). Samples were analysed in a Hellma type 1FL (1 cm path length) Macro Fluorescence cuvette (Sigma-Aldrich, Castle Hill, NSW, Australia). Absorbance spectra (240–700 nm) and EEMs were recorded with data acquisition undertaken

using Origin software (version 8.6, OriginLab Corporation, Massachusetts, USA). Wine colour measurements comprising CIELab, hue, and intensity were also recorded. Pre-processing of excitation-emission matrix (EEM) data involved normalisation according to the water Raman scattering units for the specified emission conditions, and correcting for the influence of inner filter effects (IFE), solvent background, dark detector signals, and Rayleigh masking to eliminate spectral distortion (Gilmore et al., 2017).

5. Sample preparation and A-TEEM analysis for blending experiment

Wines were added into 12 mL glass vials with Teflon lined caps to prepare the blends as shown in Table 1 to obtain a final volume of 10 mL. After addition, vials were mixed thoroughly for 60 s using a benchtop vortex and samples were prepared and analysed in duplicate as described in Section 4, but using a dilution of 1:150.

Table 1. Percentages of wine in blends of Shiraz with Cabernet Sauvignon or Grenache.

Variety	Blending percentage (v/v)							
Shiraz	100	99	95	90	85	60	50	0
Cabernet Sauvignon or Grenache	0	1	5	10	15	40	50	100

6. Statistical analysis

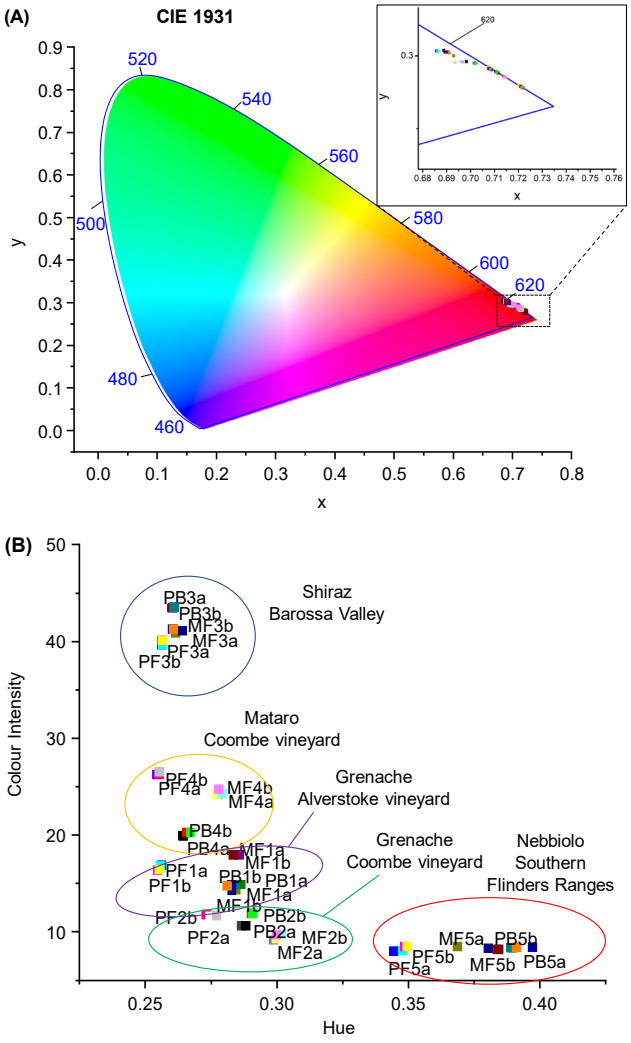
One-way analysis of variance (ANOVA) with Tukey's honestly significant difference (HSD) post hoc test for pairwise comparisons ($\alpha = 0.05$) for basic chemical measures and wine colour parameters according to stage of winemaking and region were undertaken with XLSTAT (version 2019.03.02, Addinsoft, Boston, USA). EEM data were unfolded into a two-way array using transform unfold multiway (mode 1) in Solo software (version 8.7.1, Eigenvector Research, Inc., Manson, WA, USA). Principal component analysis (PCA) was carried out with singular value decomposition and autoscale pre-processing with four principal components to explore variations in samples at different stages of winemaking using Solo software. Samples were labelled with their variety according to winemaking stage and classified using extreme gradient boosting discriminant analysis (XGBDA) after partial least squares (PLS) compression using five latent variables (LV), with mean centring pre-processing and decluttering with generalised least squares weighting (GLSW) at 0.2 to both calibrate and cross-validate (k =10, Venetian blinds procedure). The model was evaluated using confusion matrix score probabilities according to previous studies (Ranaweera, Gilmore, et al., 2021a, 2021b). For the blending experiment, unfolded EEM data were modelled with PLS and XGB regression algorithms (Solo software) using blending percentage as the y-block. Root mean square error of cross-validation (RMSECV) (Venetian blinds with 10 splits) and coefficients of

158 determination for both calibration and cross-validation (R^2 cal, R^2 CV) were used to evaluate
159 the effectiveness of the models.

160 ## RESULTS AND DISCUSSION

161 **1. Variations according to stage of winemaking**

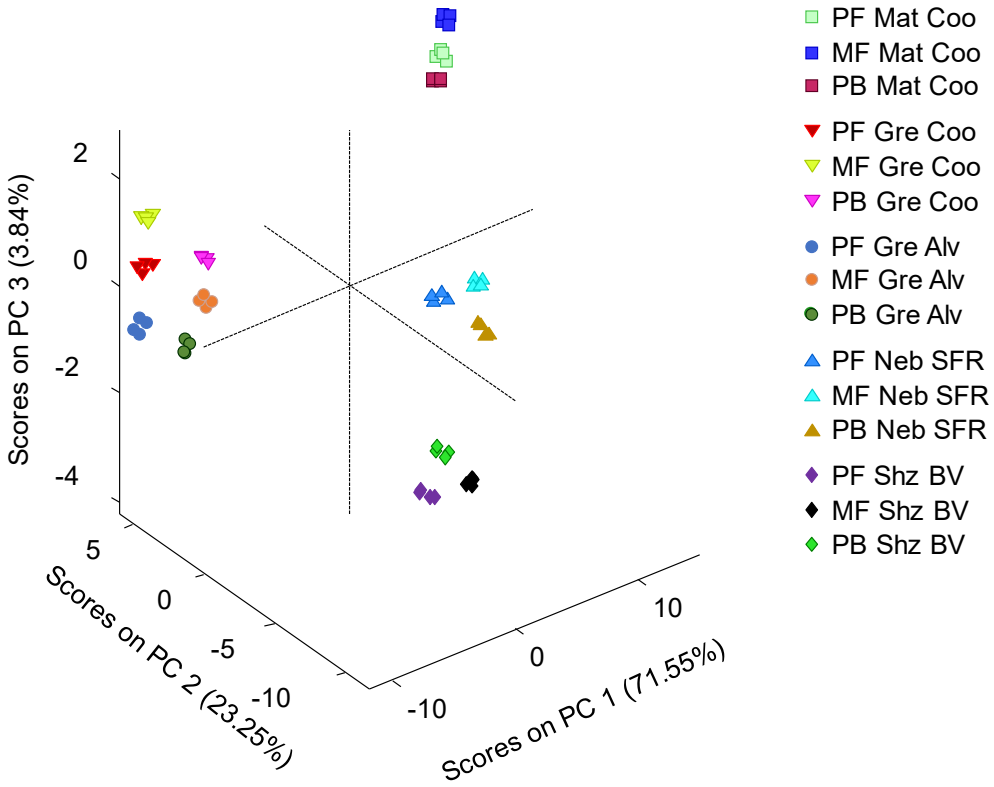
162 CIELab colour parameters and basic oenological measurements of wine samples obtained
163 during the winemaking process were assessed with one-way ANOVA according to different
164 winemaking stages as well as according to origin (for different varieties), as shown in Table S1
165 and Table S2, respectively, of the Supplementary data. When analysed according to the
166 winemaking stage (Table S1), there were no significant differences (p-value > 0.26) in basic
167 chemistry (alcohol, pH, TA) nor colour parameters (hue, intensity, L*, a*, b*, C*). Values for
168 the chromatic characteristics at the different winemaking stages showed that the wines were
169 relatively low in lightness (L*), moderately high in red (a*) and yellow (b*) , and high in
170 chroma (C*). These results generally aligned with variations among oenological properties and
171 colour expression during winemaking (Arcena et al., 2020), depending on the stage/time period
172 of sampling. According to the origin of the samples (Table S2), alcohol % v/v and all colour
173 parameters showed significant variation (p-value <0.0001), whereas pH and TA were not
174 significantly different.
175 In the CIE 1931 xyY colour space, all samples were congregated together in the red zone (x =
176 0.68 to 0.72 and y = 0.27 to 0.31, Figure 1A), which contrasted with the hue vs intensity plot,
177 where clear separation of Shiraz from Barossa Valley and Nebbiolo from Southern Flinders
178 Ranges could be observed (Figure 1B). Furthermore, Grenache and Mataro samples from
179 vineyards at Waite campus (Alverstoke and Coombe) were clustered relatively close, but were
180 still somewhat differentiated. Based on this simple analysis, it appeared that unique information
181 related particularly to origin that was not impacted by the stage of processing could be
182 expressed from absorbance data.



185 **FIGURE 1. Analysis of colour parameters according to stage of the winemaking process**
186 **(n = 15, duplicate samples analysed twice) showing (A) CIE 1931 plot (inset shows**
187 **clustering of samples) and (B) hue vs colour intensity graph. PF, post primary**
188 **fermentation; MF, post malolactic fermentation, PB, pre-blending. 1, Grenache from**
189 **Alverstoke vineyard; 2, Grenache from Coombe vineyard; 3, Shiraz from Barossa Valley;**
190 **4, Mataro from Coombe vineyard; 5, Nebbiolo from Southern Flinders Ranges.**

191 The observations were interesting but the stages of winemaking were seemingly overshadowed.
192 As such, further exploratory analysis was carried out with EEM data (which can be considered
193 as a molecular fingerprint (Gilmore et al., 2017)) using PCA (Figure 2). The first three principal
194 components explained 94.80 % of the total variance for the samples, which were perfectly
195 clustered according to origin for the different winemaking stages. Wines from the different

196 regions and varieties were reasonably well separated along PC1 except for Mataro from
197 Coombe vineyard and Shiraz from Barossa Valley. PC2 especially segregated Nebbiolo from
198 Southern Flinders Ranges from the remainder and the Shiraz to an extent, whereas Waite
199 Campus vineyard samples (Grenache and Mataro) overlapped. The Shiraz was well separated
200 from other samples along PC3. This outcome provided the first indication that the fluorescence
201 molecular fingerprint according to origin could be traced (and seemingly preserved) during
202 winemaking.



203

204 **FIGURE 2. Scores from PCA of EEM data for samples of different variety/origin collected**
205 **at three stages of winemaking (n = 15, duplicate samples analysed twice). Gre Alv,**
206 **Grenache from Alverstoke vineyard; Gre Coo, Grenache from Coombe vineyard; Mat**
207 **Coo, Mataro from Coombe vineyard; Neb SFR, Nebbiolo from Southern Flinders Ranges;**
208 **Shz BV, Shiraz, from Barossa Valley. PF, post primary fermentation; MF, post malolactic**
209 **fermentation; PB, pre-blending.**

210 XGBDA was subsequently carried out as reported in (Ranaweera, Gilmore, et al., 2021a,
211 2021b) for classification by origin. Figure 3 shows the class cross-validation (CV, Venetian

blinds CV method) prediction probability from this machine learning approach, revealing the probability of each sample belonging to the class it most closely resembles. Class CV prediction demonstrated excellent separation of samples according to origin, grouping all stages of winemaking (i.e., post primary fermentation, post malolactic fermentation, and pre-blending) together for each class. These results further emphasised the distinct possibility of tracing samples through different stages of winemaking according to their origin. Thus, EEM data from the A-TEEM technique could provide an original spectral fingerprint of the product that can be maintained during wine production, thereby opening up avenues for this being used as a chemical signature for traceability.

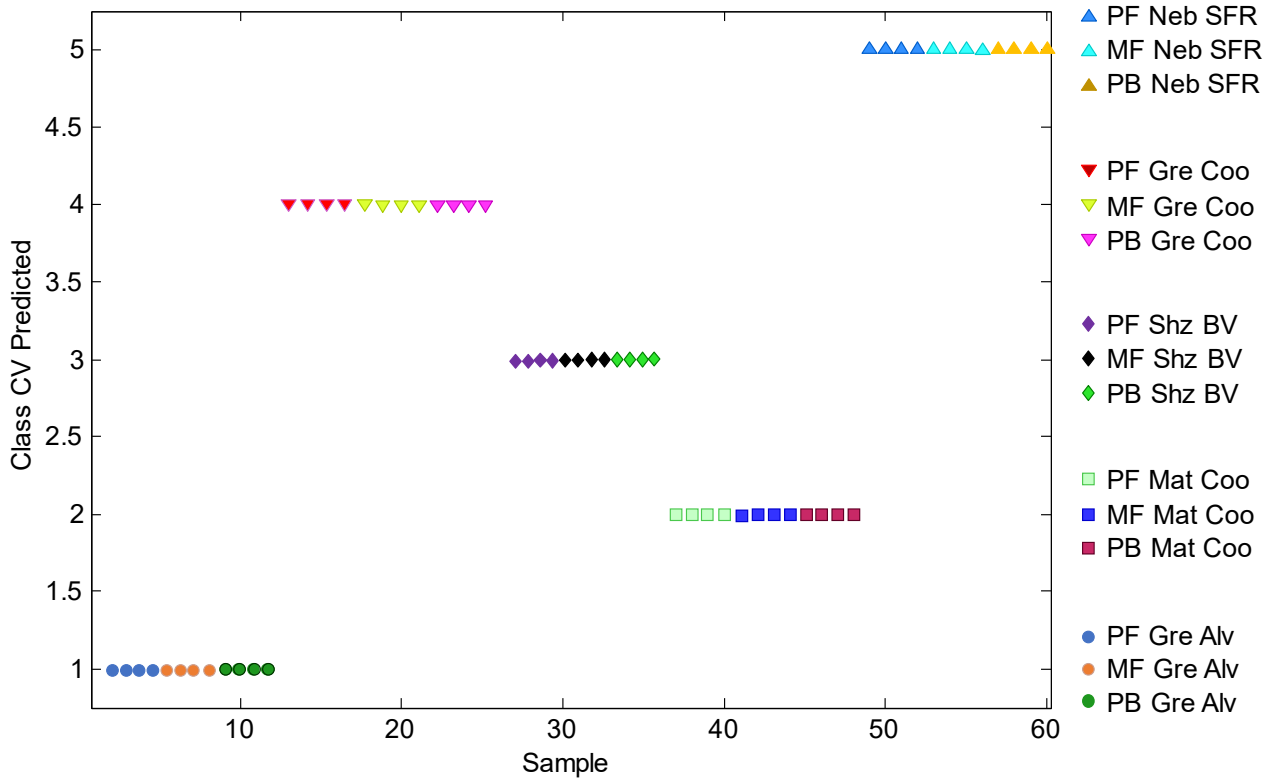


FIGURE 3. Class CV predicted for wine origin from XGBDA analysis of EEM data for samples collected at three stages of winemaking (n = 15, duplicate samples analysed twice). Gre Alv, Grenache from Alverstoke vineyard; Gre Coo, Grenache from Coombe vineyard; Mat Coo, Mataro from Coombe vineyard; Neb SFR, Nebbiolo from Southern Flinders Ranges; Shz BV, Shiraz, from Barossa Valley. PF, post primary fermentation; MF, post malolactic fermentation; PB, pre-blending.

2. Modelling to identify blend proportions

Testing the A-TEEM approach for sensitivity in terms of changes in matrix from introducing a blending component was another important consideration in terms of possible fraud detection.

To evaluate the possibility of identifying the blending percentage of each sample, regression methods were applied to EEM data for Shiraz wine containing proportions of Cabernet Sauvignon or Grenache. As a common method, PLS regression (PLSR) was applied for the two sets of wines blended according to the amounts in Table 1. The correlation between the actual blends and predicted percentages were evaluated, with R^2 CV and RMSECV values for Shiraz and Cabernet Sauvignon blends (0.996, 2.17) and Shiraz and Grenache blends (0.992, 3.12) as shown in Figure S1. Accuracy of the models was good, with R^2 CV values > 0.990 for both sets of blends, but the RMSECV values were slightly high, at 2-3 %. PLSR uses latent variables (components) that explain as much of the covariance as possible between a set of predictor X-variables and response Y-variables (Ghanem et al., 2015). A study by Gilmore et al. (2020) identified that XGB regression (XGBR) yielded more precise fits for prediction of phenolic and anthocyanin compound concentrations from A-TEEM data compared to PLSR. Therefore, XGB regression was applied to the blending experiment data to seek improvements in the regression models. Figure 4 shows the results, with the XGBR models having a perfect R^2 CV of 1.00 and exceedingly low RMSECV of 0.00028 for both sets of Shiraz blends.

XGBR can clearly predict the blend percentage for each sample, notably with a clear distinction between 0 % blend and 1 % blend for both Shiraz/Cabernet Sauvignon and Shiraz/Grenache. This was a striking result, highlighting that XGBR modelling of EEM data could be a successful option for detecting the addition of small proportions of different varietal wines. With further development and ultimately the production of databases, it is conceivable that this approach could be applied to robustly predict the composition of unknown sample blends. In addition, the approach is simple and rapid in comparison to sensitive DNA techniques (e.g., based on cultivar genotype to determine wine blends), which suffer from reproducibility problems when authenticating experimental or commercial wines (Boccacci et al., 2020).

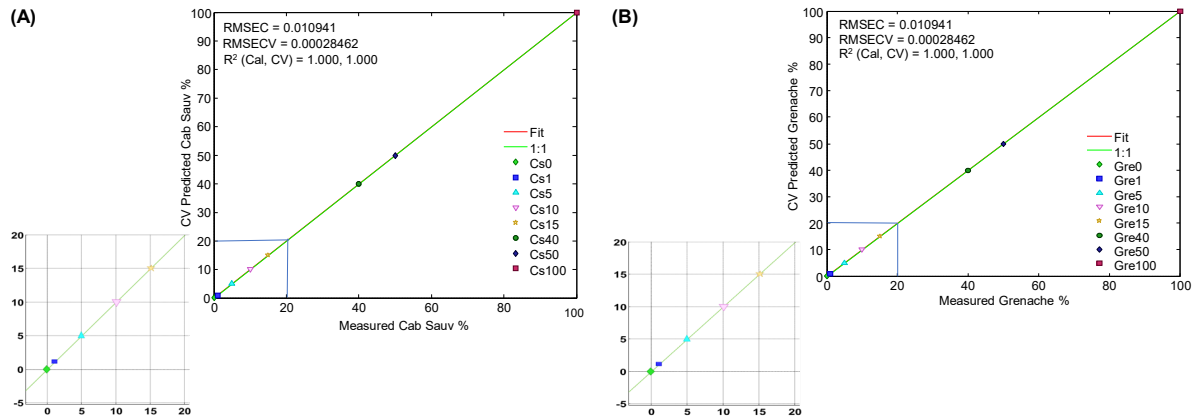


FIGURE 4. XGB regression of measured vs CV predicted blending percentages for Shiraz wine containing proportions of (A) Cabernet Sauvignon or (B) Grenache. Insets show more detail of the sample separation for 0 %–20 % blends.

CONCLUSIONS

The A-TEEM approach with machine learning modelling continued to show promise as an indispensable tool for wine authentication. In this preliminary work, A-TEEM was applied to monovarietal unfinished wine samples collected from different stages of the winemaking process (i.e., post primary fermentation, post malolactic fermentation, and pre-blending) to investigate the possibility of tracing molecular fingerprints during wine production. PCA was able to separate samples from different origins based on EEM data and subsequent XGBDA modelling could differentiate the samples with 100 % accuracy. Further highlighting the power of the A-TEEM technique, two sets of wine blends (Shiraz/Cabernet Sauvignon and Shiraz/ Grenache) were analysed to model the proportions of wine in the blend (beginning as low as 1 %). Regression models built with PLSR and XGBR were evaluated in terms of correlation coefficient and cross-validation error, with unrivalled accuracy achieved for the XGBR model with R^2 CV equal to 1.00 and small RMSECV for both sets of wine blends. Given the possibility of tracing a wine's origin through production in conjunction with identifying small additions of other wine in a blend, this approach could foreseeably be developed into a robust method and applied in the industry not only for validating the origin of wine but also detecting other aspects of wine fraud.

Acknowledgements

We are grateful to Peter Leske, Director of Revenir Winemaking, and Paul Grbin, Head Winemaker in the School of Agriculture, Food and Wine, the University of Adelaide, for supplying wine samples. We thank Lara Westaway, the University of Adelaide, for sample collection. R.K.R.R. is appreciative of the financial support provided by an Australian

Government Scholarship and Wine Australia supplementary scholarship (WA Ph1909). The Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au; project number IC170100008) is funded by the Australian Government with additional support from Wine Australia, Waite Research Institute, and industry partners. The University of Adelaide is a member of the Wine Innovation Cluster.

REFERENCES

Aceto, M., Robotti, E., Oddone, M., Baldizzone, M., Bonifacino, G., Bezzo, G., Di Stefano, R., Gosetti, F., Mazzucco, E., Manfredi, M., & Marengo, E. (2013). A traceability study on the Moscato wine chain. *Food Chemistry*, *138*(2–3), 1914-1922. https://doi.org/http://dx.doi.org/10.1016/j.foodchem.2012.11.019

Almeida, C. M. R., & Vasconcelos, M. T. S. D. (2004). Does the winemaking process influence the wine 87Sr/86Sr? A case study. *Food Chemistry*, *85*(1), 7-12. https://doi.org/https://doi.org/10.1016/j.foodchem.2003.05.003

Arcena, M. R., Kebede, B., Leong, S. Y., Silcock, P., & Oey, I. (2020). Feasibility of using integrated fingerprinting, profiling and chemometrics approach to understand (bio) chemical changes throughout commercial red winemaking: A case study on Merlot. *Food Research International*, *127*, 108767. https://doi.org/https://doi.org/10.1016/j.foodres.2019.108767

Boccacci, P., Chitarra, W., Schneider, A., Rolle, L., & Gambino, G. (2020). Single-nucleotide polymorphism (SNP) genotyping assays for the varietal authentication of 'Nebbiolo' musts and wines. *Food Chemistry*, *312*, 126100. https://doi.org/https://doi.org/10.1016/j.foodchem.2019.126100

Castiñeira, M. d. M., Brandt, R., Jakubowski, N., & Andersson, J. T. (2004). Changes of the metal composition in German white wines through the winemaking process. a study of 63 elements by inductively coupled plasma−mass spectrometry. *Journal of Agricultural and Food Chemistry*, *52*(10), 2953-2961. https://doi.org/10.1021/jf035119g

Catalano, V., Moreno-Sanz, P., Lorenzi, S., & Grando, M. S. (2016). Experimental review of DNA-based methods for wine traceability and development of a single-nucleotide polymorphism (SNP) genotyping assay for quantitative varietal authentication. *Journal of Agricultural and Food Chemistry*, *64*(37), 6969-6984. https://doi.org/10.1021/acs.jafc.6b02560

Dooley, L. M., Threlfall, R. T., Meullenet, J.-F., & Howard, L. R. (2012). Compositional and sensory impacts from blending red wine varietals. *American Journal of Enology and Viticulture*, *63*(2), 241-250. https://doi.org/10.5344/ajev.2012.11086

Ghanem, E., Hopfer, H., Navarro, A., Ritzer, M. S., Mahmood, L., Fredell, M., Cubley, A., Bolen, J., Fattah, R., Teasdale, K., Lieu, L., Chua, T., Marini, F., Heymann, H., & Anslyn, E. V. (2015). Predicting the composition of red wine blends using an array of multicomponent peptide-based sensors. *Molecules*, *20*(5), 9170-9182. http://www.mdpi.com/1420-3049/20/5/9170

Gilmore, A. (2014). How to collect national institute of standards and technology (NIST) traceable fluorescence excitation and emission spectra. In Y. Engelborghs & A. J. W. G. Visser (Eds.), *Fluorescence Spectroscopy and Microscopy: Methods and Protocols* (pp. 3-27). Humana Press. https://doi.org/10.1007/978-1-62703-649-8_1

Gilmore, A., Akaji, S., & Csatorday, K. (2017). Spectroscopic analysis of red wines with A-TEEM molecular fingerprinting. *Readout*, *2*(49), 41-48. https://www.horiba.com/in/publications/readout/article/spectroscopic-analysis-of-red-wines-with-a-teem-molecular-fingerprinting-53035/

Gilmore, A., Schober, D., Penichet, I., Zincker, J., & González, A. (2020). Quality assessment of Cabernet Sauvignon from Chile based on simultaneous absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) analysis. SciX Conference SciX Virtual Event Platform.

Imparato, G., Paolo, E. D., Braca, A., & Lamanna, R. (2011). Nuclear magnetic resonance profiling of wine blends. *Journal of Agricultural and Food Chemistry*, *59*(9), 4429-4434. https://doi.org/10.1021/jf200587n

Jones-Moore, H. R., Jelley, R. E., Marangon, M., & Fedrizzi, B. (2022). The interactions of wine polysaccharides with aroma compounds, tannins, and proteins, and their importance to winemaking. *Food Hydrocolloids*, *123*, 107150. https://doi.org/https://doi.org/10.1016/j.foodhyd.2021.107150

Li, S.-Y., Zhao, P.-R., Ling, M.-Q., Qi, M.-Y., García-Estévez, I., Escribano-Bailón, M. T., Chen, X.-J., Shi, Y., & Duan, C.-Q. (2020). Blending strategies for wine color modification I: Color improvement by blending wines of different phenolic profiles testified under extreme oxygen exposures. *Food Research International*, *130*, 108885. https://doi.org/https://doi.org/10.1016/j.foodres.2019.108885

Lonvaud-Funel, A. (2010). Effects of malolactic fermentation on wine quality. In A. G. Reynolds (Ed.), *Managing Wine Quality* (pp. 60-92). Woodhead Publishing. https://doi.org/https://doi.org/10.1533/9781845699987.1.60

Preserova, J., Ranc, V., Milde, D., Kubistova, V., & Stavek, J. (2015). Study of phenolic profile and antioxidant activity in selected Moravian wines during winemaking process by FT-IR spectroscopy. *Journal of Food Science and Technology*, *52*(10), 6405-6414. https://doi.org/10.1007/s13197-014-1644-8

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021a). Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. *Food Chemistry*, *335*, 127592. https://doi.org/https://doi.org/10.1016/j.foodchem.2020.127592

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021b). Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. *Food Chemistry*, *361*, 130149. https://doi.org/https://doi.org/10.1016/j.foodchem.2021.130149

Ranaweera, R. K. R., Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Authenticity and traceability in the wine industry: From analytical chemistry to consumer perceptions. In G. Smithers (Ed.), *Reference Module in Food Science*. Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-08-100596-5.22876-X

Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021). Defining wine typicity: Sensory characterisation and consumer perspectives. *Australian Journal of Grape and Wine Research*, *27*(2), 246-256. https://doi.org/https://doi.org/10.1111/ajgw.12474

369 Styger, G., Prior, B., & Bauer, F. (2011). Wine flavor and aroma. *Journal of Industrial*
370 *Microbiology & Biotechnology*, *38*(9), 1145-1159. https://doi.org/10.1007/s10295-011-
371 1018-4

372 Versari, A., Laurie, V. F., Ricci, A., Laghi, L., & Parpinello, G. P. (2014). Progress in
373 authentication, typification and traceability of grapes and wines by chemometric
374 approaches. *Food Research International*, *60*, 2-18.
375 https://doi.org/https://doi.org/10.1016/j.foodres.2014.02.007

376 Wine Australia. (2017). *The history, evolution and revolution of Australian red blends*.
377 Retrieved 05.10.2021 from https://www.wineaustralia.com/whats-happening/stories-
378 of-australian-wine/july-2017/red-blends

379 Wine Australia. (2018). *The blending rules*. Retrieved 2/10/2018 from
380 https://www.wineaustralia.com/labelling/further-information/the-blending-rules

381

OENO One | By IVES International Viticulture & Enology Society
www.ives-openscience.eu

SUPPLEMENTARY DATA

Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines

TABLE S1. Least squares means and one-way ANOVA results for oenological parameters and colour measurements (L*, lightness; a*, green–red; b*, blue–yellow; C*, chroma) of wine samples according to stage of winemaking.

Stage of winemaking[1]	Alcohol (% v/v)	pH	Titratable acidity (g/L	Hue	Intensity	L*	a*	b*	C*
PF	14.4	3.50	5.7	0.28	20.6	29.6	71.8	65.7	97.4
MF	14.5	3.60	5.0	0.30	20.5	29.4	71.4	64.5	96.3
PB	14.5	3.43	5.5	0.30	19.4	29.5	71.7	65.6	97.3
p-value	0.945	0.417	0.261	0.710	0.988	0.999	0.996	0.938	0.979

[1]PF, post primary fermentation; MF, post malolactic fermentation; PB, pre-blending

OENO One | By IVES International Viticulture & Enology Society
www.ives-openscience.eu

SUPPLEMENTARY DATA

Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines

TABLE S2. Least squares means and one-way ANOVA results for oenological parameters and colour measurements (L*, lightness; a*, green–red; b*, blue–yellow; C*, chroma) of wine samples according to origin for different varieties. Different letters within a column indicate statistically significant differences among the means according to Tukey HSD post hoc test (α = 0.05).

Sample origin[1]	Alcohol (% v/v)	pH	Titratable acidity (g/L)	Hue	Intensity	L*	a*	b*	C*
NebSF	13.9 c	3.42	6.3	0.37 a	8.3 d	37.4 a	78.1 a	65.3 b	101.8 ab
GreCoo	15.3 a	3.60	5.0	0.29 b	10.8 d	34.1 ab	76.8 a	64.3 b	100.2 ab
GreAlv	14.6 b	3.30	5.1	0.27 b	16.5 c	31.5 b	75.5 a	72. 0 a	104.3 a
MatCoo	14.7 b	3.58	5.7	0.27 b	23.6 b	26.1 c	68.8 b	67.9 ab	96.6 b
ShzBV	14.0 c	3.62	5.0	0.26 b	41.5 a	18.5 d	59.1 c	56.8 c	82.0 c
p-value	<0.0001	0.073	0.073	<0.0001	<0.0001	<0.0001	<0.0001	0.000	<0.0001
Significant	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes

[1]NebSF, Nebbiolo from Southern Flinders Ranges; GreCoo, Grenache from Coombe vineyard at Waite campus; GreAlv, Grenache from Alverstoke vineyard at Waite campus; MatCoo, Mataro from Coombe vineyard at Waite campus; ShzBV, Shiraz from Barossa Valley.

OENO One | By IVES International Viticulture & Enology Society
www.ives-openscience.eu

SUPPLEMENTARY DATA

Feasibility of using spectrofluorometric analysis to trace wine molecular fingerprint through the winemaking process and recognise the blending percentage of different varietal wines

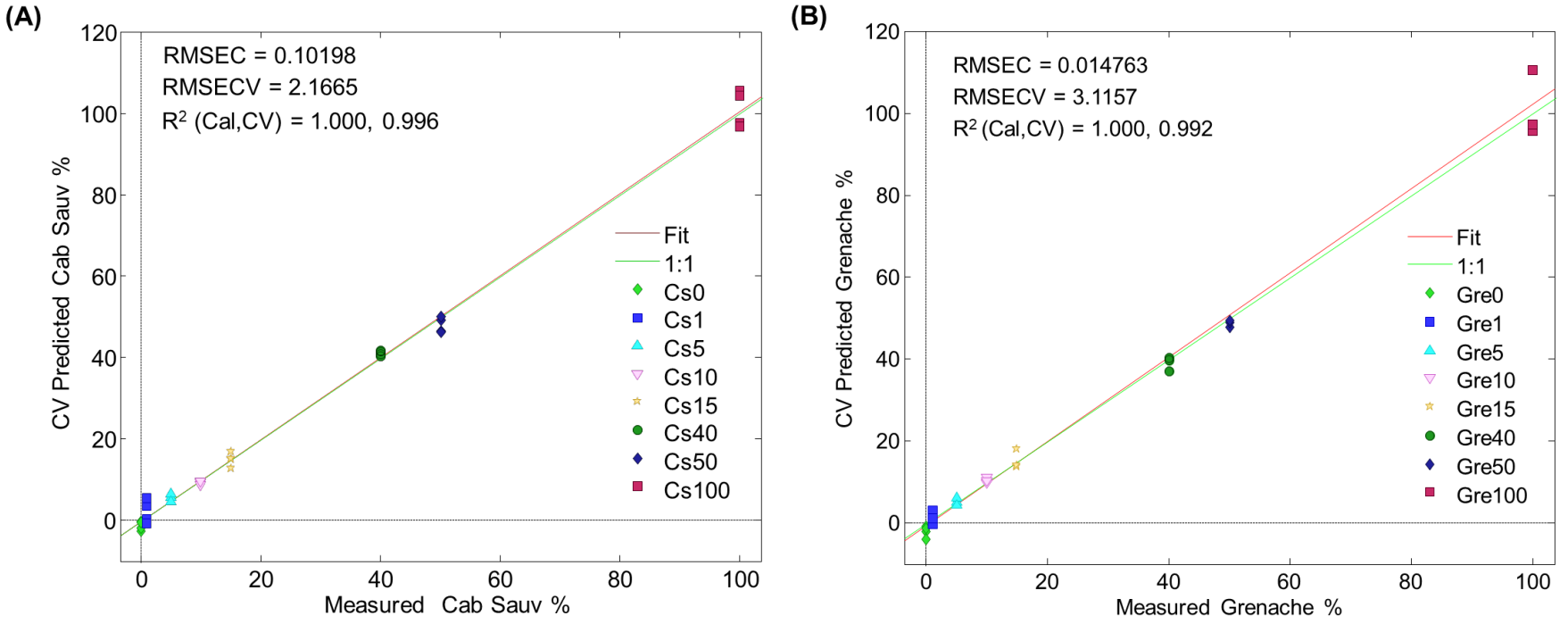


FIGURE S1. PLS regression of measured vs CV predicted blending percentages for Shiraz wine containing proportions of (A) Cabernet Sauvignon and (B) Grenache.

Chapter 7

Concluding Remarks & Future Perspectives

7.1 Conclusions

Wine authentication is an extensive concept that includes verifying wine age, geographical origin, variety, and production practices to be compliant with label descriptions and other regulatory requirements. Perhaps more evident recently, many of the popular wines of the world are every so often the subject of substitution or counterfeiting, with no exception for wines from Australia. The ability to verify important parameters encompassed within the chemical composition of wines would be beneficial in making an Australian wine unique in the market among others. Upon reviewing the literature as presented in **Chapters 1** and **2,** it was identified that different analytical techniques have been developed to verify wine authenticity over the past few decades, but considering the multifaceted production process, it has remained a challenge to establish a robust method to authenticate the inherent attributes of a wine. This thesis aimed to address this issue by developing a robust but simple analytical method to accurately authenticate wine according to geographical origin, variety and vintage. In a novel approach to wine authentication, the use of absorbance and fluorescence spectroscopy in conjunction with machine learning has been assessed. In addition, the work has addressed the underlying chemical markers attributable to regional wines and has explored models for predicting chemical parameters using spectral data.

7.1.1 Development of methodology using fluorescence spectroscopy for wine authentication

Fulfilling Objective 1 of the project, spectrofluorometric analysis based on the A-TEEM technique was initially applied to authenticate the geographical origin of Australian Cabernet Sauvignon wine from three Geographical Indications (GI) – Coonawarra, Yarra Valley, and Margaret River – together with wines from Bordeaux, France (**Chapter 3**). The A-TEEM technique was applied as a non-targeted method that uses a three-dimensional excitation-emission matrix (EEM) to obtain a "molecular fingerprint" from fluorophoric components present in wine to differentiate them according to region of origin. The EEM data were pre-processed prior to chemometric

analysis by normalising according to the water Raman scattering units and correcting for the influence of inner filter and other effects to eliminate spectral distortion. Classification accuracy according to A-TEEM was compared to models developed with element concentrations determined by ICP-MS as a reference method for wine authentication. ICP-MS data modelling was undertaken with discriminant analysis (DA) and the EEM data were analysed using support vector machine discriminant analysis (SVMDA). Additionally, a novel application of machine learning using extreme gradient boosting discriminant analysis (XGBDA) was used for element and EEM data. The effectiveness of the methods was evaluated by cross-validation of classification models, using evaluation parameters including sensitivity, specificity, precision and F1 score. DA of elemental data and the SVMDA of EEM data resulted in an overall correct classification of 94.2 % and 84.7 %, respectively. In comparison, the analysis of EEM data with XGBDA afforded 100 % correct classification for all classes whereas ICP-MS resulted in an overall 97.7 % correct classification with XGBDA. From this initial study it was concluded that the A-TEEM technique was a simple, rapid, and sensitive method and modelling of EEM data was highly effective in discrimination of wine according to geographical origin. The unrivalled results of this investigation led to further expansion of this approach to other varieties and regions, and exploration of the underlying chemical markers leading to the classification.

7.1.2 Verification of the method according to region and exploring the varietal authentication with insight into chemical markers

In accordance with Objective 2 of the overall study, the method developed in **Chapter 3** was further tested with a broad range of commercial, unreleased wines (n = 221) from ten GI in South Australia and Western Australia belonging to three different varieties (Shiraz, Cabernet Sauvignon, and Merlot), as reported in **Chapter 4**. Notably, EEM and absorbance data were combined through multi-block data analysis to improve the effectiveness of the machine learning classification. The results reinforced the capability of A-TEEM technique in combination with XGBDA for wine authentication, providing 100 % accurate classification of wines for variety

and 99.7 % for region of origin. As for the aim of understanding the chemicalmarkers that may discriminate among wines, the main fluorophoric components attributed to variation in the samples were proposed using parallel factor analysis (PARAFAC) of EEM data. PARAFAC tentatively revealed the most dominant fluorescent compounds; however those components did not contribute to the clustering of samples related to wine origin due to overfitting of all the classes with the same loading. Considering the potential utility of the methodology, A-TEEM data were applied to predict the concentrations of 24 phenolic compounds in wine based on HPLC reference data and partial least squares regression. The overall results afforded values for Pearson's correlation (r) > 0.99 and coefficient of determination (adjusted R^2) > 0.99, with standard error of the regression slope < 0.002, thus indicating the high explanatory power and robustness of the prediction models. Even though analysis of additional samples was recommended to strengthen the models, the A-TEEM technique with chemometrics was revealed to be a rapid method for phenolic quantification as well as a valuable tool for authentication of wine. Overall, a potential insight into the underlying chemistry of the wines was provided, which may ultimately be used to improve understanding of wine regionality.

7.1.3 Investigation of intraregional variation of Shiraz wines using A-TEEM and machine learning modelling

The A-TEEM technique was applied for the first time to examine intraregional differences using a set consisting primarily of experimental Shiraz wines (with some unreleased commercial wines) from five proposed subregions (not delimited) of the Barossa Valley GI along with Eden Valley GI (**Chapter 5**). The molecular fingerprint of wine was captured using the A-TEEM technique to reveal variations ascribed to the terroir within the same region, addressing Objective 3 of the project. The wines came from three consecutive seasons (2018, 2019 and 2020), with the experimental wines being produced in a consistent manner, permitting authentication testing according to the vintage as well. Using fusion of A-TEEM data in a multi-block approach, initial exploratory analyses involved k-means clustering and PCA, with k-means analysis revealing clustering according to the vintage year, emphasising

the strong seasonal impacts for wine, whereas PCA plots of individual years showed substantial clustering of some subregions. Clustering of subregions by PCA was especially evident in 2018, whereas environmental factors may have obscured the differentiation in the subsequent vintages. Supervised classification with XGBDA models afforded an unsurpassable level of accuracy (i.e., 100 %) according to vintage year as well as for subregion across all vintages. The A-TEEM method was shown to be a powerful tool for authenticating wine from origins that are in close proximity and results of this study may provide impetus for employing this as a chemical-based approach to better understand the expression of terroir in regional Australian wines. In another novel application of A-TEEM data, predictive models were also developed in **Chapter 5** for basic chemical parameters by regressing against data obtained from reference methods. Thus, XGB regression with cross-validation afforded strong correlations (R^2 CV) and low errors (RMSECV) for alcohol (0.877, 0.22 % v/v) glucose + fructose (0.888, 0.12 g/L), pH (0.793, 0.06), titratable acidity (0.839, 0.28 g/L), and volatile acidity (0.876, 0.06 g/L), showing the capability of A-TEEM and chemometrics as an approach for rapid determination of these useful chemical parameters.

7.1.4 Tracing the molecular fingerprint of wine throughout the winemaking process and recognising the blending percentage of varietal wine

Chapter 6 describes a preliminary investigation of the impact of production practices on wine authentication using the A-TEEM approach, with monovarietal wines from three different processing stages (post primary fermentation, post malolactic fermentation, and pre-blending). This study provided a first evaluation of the stability of the molecular fingerprint of wine throughout the winemaking process and offered insight into whether A-TEEM could be used for traceability at different stages of production. A plot of hue and colour intensity obtained from absorbance data of the samples as well as PCA of the EEM data showed a clustering of wine samples according to their origin, thus showing the lack of

influence of stage of processing and highlighting the prospect of tracing a wine's spectral signature through production. Similar to the statistical results of previous experiments, XGBDA showed a remarkable ability to differentiate samples according to their origin with 100% accuracy irrespective of winemaking stage, highlighting the capability of the approach based on the selectivity of EEM data. For the blending experiment, unreleased but commercially-produced monovarietal wines were prepared in two sets, Shiraz containing Cabernet Sauvignon and Shiraz containing Grenache, with a series of blends comprising 1 %, 5 %, 10 %, 15 %, 20 %, 40 %, and 50 % of the second varietal wine. Regression methods were applied to EEM data in an effort to identify the blending percentage. Comparing two regression methods, PLSR and XGBR with both sets of blends, XGBR was found to yield more precise prediction of percentage in the blend, achieving R^2 CV of 1.00 and RMSECV of 0.00028 in comparison to PLSR (R^2 CV > 0.99 and RMSECV between 2-3). Although more work is required in this space, A-TEEM analysis with XGBR could be an excellent option for the detection of possible fraud in relation to wine blending.

In summary, this thesis has contributed to identifying a reliable method to accurately authenticate wine according to geographical origin, variety, and vintage. Application of spectrofluorometric analysis for wine authentication in combination with machine learning modelling using XGBDA provided unapparelled results and the approach could conceivably be adapted for use in a supply chain setting for validating the origin of commercial wine. This project also contributed knowledge regarding the influence of terroir on wine composition, as evidenced by variations in wine molecular fingerprints. This is important when trying to determine regional and varietal characters encompassed within the chemical composition of wines, and provides the opportunity for optimising perceived value in the global market. Finally, the assessment of processing stages during winemaking and of wine blending is a step forward in terms of a using wine chemistry within a traceability system to aid in mitigating wine fraud. Ultimately, the outcomes of this thesis have enhanced scientific knowledge regarding methodology for wine authentication and provided the foundation for commercial application of this approach in the wine industry.

7.2 Future perspectives

The possibility of authentication of wine according to its geographical origin, variety, and vintage using fluorescence spectroscopy as a rapid and sensitive method has been successfully revealed in **Chapters 3**, **4**, and **5**. Specifically, according to **Chapter 5,** discrimination of wine at a subregional level was possible due to the sensitivity of the technique and capability of machine learning models. This raises the possibility of using the approach developed in this thesis as a chemical basis to verify regional zoning based on terroir or to authenticate individual vineyards or estates, although these aspects would need to be verified. Additionally, identifying the potential of this technique to classify wine after a certain period of time based on chemometric models built previously needs to be explored, since wines can be stored in the bottle for several years or transported to different countries. Hence, the influence of storage conditions related to temperature and time on the classification of wine also needs to be investigated. Leading on from this, it would be useful to identify the applicability of the models developed for use on the same wines after ≈ 5 years of storage under different conditions.

From the prediction modelling of phenolic compound concentrations (**Chapter 4**), it was revealed that phenolic compounds could contribute significantly as chemical drivers for classification. Furthermore, verification of the chemical markers represented in the EEM that are responsible for correct classification would be beneficial. With improved understanding of chemical composition with respect to Australian wines of provenance, future work could investigate the most appropriate viticultural or winemaking practices that preserve or increase wine quality and expression of regionality based on A-TEEM analysis in conjunction with sensory assessment. Indeed, A-TEEM with machine learning modelling could conceivably be utilised during wine production to assist in the targeting of wines with desired style and quality parameters, although this remains to be explored.

As described in **Chapters 4** and **5**, the A-TEEM technique with machine learning modelling is not only applicable to wine authentication processes (classification), but

can also be applied in various analyses related to wine (regression). In accordance with the preliminary results detailed in **Chapter 6**, this technique has been able to discriminate wine at different stages of winemaking. This may be further developed for improved understanding of the impacts of viticulture and winemaking practices, like the effect of additives, filtration, different vineyard treatments, clones, and rootstocks, in the pursuit of quality wines representing regional and varietal characteristics, especially under changing environmental conditions.

Application of A-TEEM in combination with XGBDA machine learning technique using cross-validation was revealed to be a robust method for wine authentication. Specifically, the Venetian blinds (k-fold) method used in this study was easy to implement and effective in analysing a different number of data splits in comparison to leave-one-out cross-validation, which is often applied in other studies but is only recommended for small data sets. It is understood that model validation plays a major role in authentication studies, and as seen in many reports, a limitation is often related to the number of samples obtained. Ideally, samples would be partitioned into calibration, validation, and test sets, but that requires a large number of samples per class, so researchers tend to stop at cross-validation. Therefore, to strengthen the models in future, inclusion of many additional samples would allow for a separate test set to be applied to assess the performance of the model. Moreover, this could potentially be considered on an industrial scale, whereby the use of A-TEEM and development of cloud-based processing and database management could provide advantages, namely the construction of a robust authenticity database by non- specialist operators who could input data into a system that is also accessible by authorities. Furthermore, maintaining a database with ongoing collection of samples will facilitate a systematic authentication approach. Overall, building on this work through incorporation of innovative technologies with chemical traceability enhances the process of wine authentication and will ultimately result in better protection of wines from designated origins as well as enhanced consumer satisfaction.

List of Abbreviations

A-TEEM	absorbance-transmission and fluorescence excitation emission matrix
AAS	atomic absorption spectroscopy
ANOVA	analysis of variance
ATR	attenuated total reflectance
ANN	artificial neural network
CA	cluster analysis
CART	classification and regression tree
CDA	canonical discriminant analysis
CP	conducting polymers
CV	cross-validation
DA	discriminant analysis
DART	direct analysis in real time
DT	decision tree
EEM	excitation-emission matrix
EM	electromagnetic
FID	flame ionisation detector
FT	Fourier transform
FTICR-MS	Fourier transform ion cyclotron resonance-mass spectrometry
GC	gas chromatography
GDD	growing degree days
GI	Geographical Indication
GLSW	generalised least squares weighting
GS1	Global Standards One
HCA	hierarchical cluster analysis
HPLC	high performance liquid chromatography
HS	static headspace
HSD	honestly significant difference
ICP-MS	inductively coupled plasma-mass spectrometry
IFE	inner filter effects
IR	infrared
IRMS	isotope ratio mass spectrometry
LC	liquid chromatography
LDA	linear discriminant analysis
LLE	liquid-liquid extraction
LMWP	low molecular weight phenolic
LV	latent variable
MALDI-MS	matrix-assisted laser desorption ionisation mass spectrometry
MFA	multiple factor analysis
MIR	mid-infrared

MS	mass spectrometry
NCM	nearest class mean
NIR	near-infrared
NMR	nuclear magnetic resonance
OIV	International Organisation of Vine and Wine
OPLS-DA	orthogonal projections to latent structures-discriminant analysis
PARAFAC	parallel factor analysis
PCA	principal component analysis
PDA	photo diode array
PDO	Protected designation of origin
PGI	Protected geographical indication
PLSDA	partial least squares-discriminant analysis
PLSR	partial least squares regression
PTR-MS	proton transfer reaction-mass spectrometry
q-NMR	quantitative nuclear magnetic resonance
Q-TOF-MS	quadrupole time-offlight-mass spectrometry
RMSEC	root mean square error of calibration
RMSECV	root mean square error of cross-validation
SAW	surface acoustic wave sensors
SE	standard error of regression slope
SERS	surface-enhanced Raman spectroscopy
SIMCA	soft independent modelling of class analogy
SLDA	stepwise linear discrimination
SNIF-NMR	site-specific natural isotopic fractionation nuclear magnetic resonance
SNP	single nucleotide polymorphism
SPE	solid-phase extraction
SPME	solid-phase micro extraction
SSR	simple sequence repeat
SVM	support vector machine
TA	titratable acidity
TIMS	thermal ionisation mass spectrometry
UHPLC	ultra-high performance liquid chromatography
XGBDA	extreme gradient boosting discriminant analysis
XGBR	extreme gradient boosting regression

List of Figures

Chapter 1

Figure 1	Chemical composition of dry red table wine with major components given on a weight-for-weight (w/w) % basis, and "Everything else" presented in mgL^{-1} . Adapted from Waterhouse et al. (2016b).	Pg 7
Figure 2	Major organic acids present in wine from grape and microbiological sources.	Pg 7
Figure 3	Examples of phenolic compounds present in wine that are classified as flavonoids and non-flavonoids.	Pg 8
Figure 4	Major amino acids present in wine.	Pg 9
Figure 5	Trace volatiles found in wine include rotundone (left), methoxypyrazines such as IBMP (center), and polyfunctional thiols such as 3-SH (right).	Pg 9
Figure 6	Scheme of the working principles of e-nose and e-tongue. Reproduced from Rodríguez-Méndez, M.L., De Saja, J.A., González-Antón, R., García-Hernández, C., Medina-Plaza, C., García-Cabezón, C., and Martín-Pedrosa, F., 2016. Electronic noses and tongues in wine industry. Front. Bioeng.Biotechnol4:81. Copyright 2016. License at https://creativecommons.org/licenses/by/4.0/	Pg 14
Figure 7	Representation of the electromagnetic spectrum. Reproduced from Ronan P.,2007. EMspectrum.svg, https://upload.wikimedia.org/wikipedia/commons/f/f1/EM_spectrum.svg. Copyright 2007. License at https://creativecommons.org/licenses/by-sa/3.0/.	Pg 15
Figure 8	Bevscan: Through bottle beverage analyzer and classifier. Images courtesy of Jeffress Engineering, www.jeffress.com.au.	Pg 16

Figure 9	Geographical classification of Chardonnay grape samples from different regions using MFA on various volatiles, basic juice parameters (titratable acidity, pH, total soluble solids), and elements, showing scores on F1 and F2 (A) and F1 and F3 (B), and corresponding loadings (C and D).Reproduced from Gambetta, J.M., Cozzolino, D., Bastian, S.E.P., and Jeffery,D.W.,2017. Exploring the effects of geographical origin on the chemical composition and quality grading of Vitis vinifera L. cv. Chardonnay grapes. Molecules, 22:218. Copyright 2017. License at https:// creativecommons.org/licenses/by/4.0/.	Pg 17
Figure 10	Example of simple wine bottle seal and closure (A) that shows evidence of damage when the bottle is opened (B). Image courtesy of David Jeffery	Pg 20
Figure 11	Example of commercially available Bubbletag® tamper evident seal showing various features including (A) brand logo, (B) chaosmetric tag, (C) tag substrate, (D) quick response (QR) code, and (E) batch identification code. Image courtesy of Prooftag, www.prooftag.net.	Pg 20
Figure 12	Example of commercially available holographic Variogram® tamper evident seal showing various features including (A) brand logo, (B) quick response (QR) code, (C) batch identification code, (D) holographic tag, and (E) tag substrate. Image courtesy of Prooftag, www.prooftag. net.	Pg 21
Figure 13	Examples of commercially available covert and overt technologies (front of bottle) showing (A) tamper-proofing seal, (B) near field communication (NFC) tag, and (C) variable data printing. Images courtesy of All4Labels, www.all4labels.com.	Pg 22
Figure 14	Examples of commercially available covert and overt technologies (rear of bottle) showing (A) holographic label, (B) laser-coding printing, and (C) tamper evident seal, (D) RFID feature, (E) QR code. Images courtesy of All4Labels, www.all4labels.com.	Pg 22
Figure 15	Different types of barcodes: (A) EAN-13, (B) interleaved 2 of 5 (Code 25, I2of5, ITF, I25), (C) Code	Pg 24

39, (D) MaxiCode and (E) PDF-417. Reprinted from Youssef, S.M. and Salem, R.M., 2007. Automated barcode recognition for smart identification and inspection automation. Expert Syst.Appl. 33,968–977, Copyright (2019), with permission from Elsevier.

Chapter 2

Figure 1 Bibliometric map of wine science-related research visualised from 222 publications (from 1990 to 2021) recovered from Web of Science Core Collection using 'wine authentication' and 'spectroscopy' analysis and figure as keywords. Literatureconstruction were facilitated with VOSviewer [12]. Pg 42

Figure 2 The electromagnetic spectrum and its relevance to different spectroscopic methods. FIR, far-infrared; MIR, mid- infrared; NIR, near-infrared; Vis, visible; UV, ultraviolet. Techniques in this review that are applied for wine authentication are indicated in light blue font. Conceptualised from [15]. Pg 43

Figure 3 Schematic of the steps involved with chemometric analysis of spectroscopic data. Stages where data fusion can be applied depending on the extent of processing are also presented in the diagram. Pg 50

Chapter 3

Figure 1 PCA biplot showing scores for the samples from the regions and loadings for the variables (significantly different elements) and (b) stepwise (forward) DA classification of regions based on PCA data; confidence ellipses correspond to a 95% confidence level. Region abbreviations: Coonawarra (C), Margaret River (M), Yarra Valley (Y), and Bordeaux (B Pg 62

Figure 2 XGBDA analysis of significant multi-element data from ICP-MS for wines of different regions Pg 63

showing (a) class prediction probability for each region and (b) class CV predicted most probable for all regions.

Figure 3 — EEM contour maps of wine samples from (a) Bordeaux, (b) Coonawarra, (c) Margaret River, and (d) Yarra Valley, after correction for the IFE, Rayleigh masking and normalisation according to water Raman scattering. — Pg 64

Figure 4 — XGBDA analysis of EEM data for wines of different regions showing (a) class prediction probability for each region (b) class CV predicted most probable for all regions. — Pg 65

Figure S1 — XGBDA variable score plot for ICP-MS multi-element data giving a measure of a variable's importance for building the model — Pg 69

Figure S2 — SVMDA analysis of EEM data for wines of different regions showing (a) class prediction probability for each region and (b) class CV predicted for all regions. — Pg 70

Chapter 4

Figure 1 — Class CV predicted by region from XGBDA analysis of A-TEEM data for varietal red wines (n = 221, analysed in duplicate) from 10 different GIs of Australia. Only one sample (circled) was misclassified. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) — Pg 77

Figure 2 — Class CV predicted for variety from XGBDA analysis of A-TEEM data for red wines (n = 221, analysed in duplicate) from 10 Australian GIs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) — Pg 78

Figure 3 — EEM contour plots of the PARAFAC model with five components indicating the fluorescent properties of typical fluorophores in wine. Tentative assignments: C1, catechin; C — Pg 79

	2, anthocyanin; C 3, epicatechin; C 4, phenolic acids; C 5, riboflavin).	
Figure 4	PCA biplot showing scores for the wine samples according to a) variety and b) region of origin along with the loadings based on phenolic compound concentrations. Abbreviations: gluc, glucoside; gal, galactoside; ac, 3-O-acetylglucoside; coum, 3-O-coumaroylglucoside.	Pg 80
Figure S1	Locations of the GIs for the samples collected from Australian regions	Pg 88
Figure S2	Split-half analysis for (a) excitation and (b) emission of the five components in the non-negativity constrained PARAFAC model.	Pg 89
Figure S3	Fluorescence loadings for (a) excitation and (b) emission of the five components in the non-negativity constrained PARAFAC model	Pg 89

Chapter 5

Figure 1	K-means partitional cluster analysis (k = 3) of A-TEEM data for subregional Shiraz wines (n = 186) from three different vintages (Class 1: 2018 in green; Class 2: 2019 in red; Class 3: 2020 in blue).	Pg 117
Figure 2	Scores for the wine samples from PCA of multi-block data according to subregion for individual vintage years, showing a) 2018, b) 2019, and c) 2020. CG, Central Grounds (blue); EE, Eastern Edge (purple); EV, Eden Valley (green); NG, Northern Grounds (pink); SG, Southern Grounds (yellow); WR, Western Ridge (brown).	Pg 118
Figure 3	Class CV predicted with XGBDA classification using A-TEEM data for subregional Shiraz wines (n = 186) according to a) vintage year for wine produced in 2018	Pg 119

	(orange triangles), 2019 (purple triangles) and 2020 (green stars), and b) subregion within the Barossa Valley GI along with Eden Valley GI for all three vintages. CG, Central Grounds (blue triangles); EE, Eastern Edge (purple triangles); EV, Eden Valley (green stars); NG, Northern Grounds (pink circles); SG, Southern Grounds (yellow diamonds); WR, Western Ridge (brown squares)	
Figure 4	Class CV predicted by region/subregion across three vintages (2018–2020) from XGBDA modelling of A-TEEM data for Shiraz wines (n = 186) from different subregions within the Barossa GI and from Eden Valley GI, South Australia. CG, Central Grounds; EE, Eastern Edge; EV, Eden Valley; NG, Nothern Grounds; SG, Southern Grounds; WR, Western Ridge	Pg 120
Figure 5	XGB regression of measured vs predicted parameters using reference methods and A-TEEM data for subregional Shiraz wines (n = 186) showing a) alcohol (% v/v), b) glucose + fructose (g L-1), c) pH, d) titratable acidity (g L-1), and e) volatile acidity (g L-1) along with calibration and validation statistics.	Pg 121
Figure A.1	Locations of the subregions for the samples collected from the Barossa GI, South Australia.	Pg 125

Chapter 6

Figure 1	Analysis of colour parameters according to stage of the winemaking process (n = 15, duplicate samples analysed twice) showing (A) CIE 1931 plot (inset shows clustering of samples) and (B) hue vs colour intensity graph. PF, post primary fermentation; MF, post malolactic fermentation, PB, pre-blending. 1, Grenache from Alverstoke	Pg 138

	vineyard; 2, Grenache from Coombe vineyard; 3, Shiraz from Barossa Valley;4, Mataro from Coombe vineyard; 5, Nebbiolo from Southern Flinders Ranges.	
Figure 2	Scores from PCA of EEM data for samples of different variety/origin collected at three stages of winemaking (n = 15, duplicate samples analysed twice). GreAlv, Grenache from Alverstoke vineyard; GreCoo, Grenache from Coombe vineyard; MatCoo, Mataro from Coombe vineyard; NebSF, Nebbiolo from Southern Flinders Ranges; ShzBV, Shiraz, from Barossa Valley. PF, post primary fermentation; MF, post malolactic fermentation:PB, pre-blending.	Pg 139
Figure 3	Class CV predicted for wine origin from XGBDA analysis of EEM data for samples collected at three stages of winemaking (n = 15, duplicate samples analysed twice). GreAlv, Grenache from Alverstoke vineyard; GreCoo, Grenache from Coombe vineyard; MatCoo, Mataro from Coombe vineyard; NebSF, Nebbiolo from Southern Flinders Ranges; ShzBV, Shiraz, from Barossa Valley. PF, post primary fermentation; MF, post malolactic fermentation:PB, pre-blending.	Pg140
Figure 4	XGB regression of measured vs CV predicted blending percentages for Shiraz wine containing proportions of (A) Cabernet Sauvignon and (B) Grenache. Insets show more detail of the sample separation for 0%–20% blends.	Pg 142
Figure S1	PLS regression of measured vs CV predicted blending percentages for Shiraz wine containing proportions of (A) Cabernet Sauvignon and (B) Grenache.	Pg 148

List of Tables

Chapter 1

Table 1	Examples of mass spectrometry approaches used for wine authentication	Pg 12
Table 2	Examples of separation and detection techniques used for wine authentication	Pg 12
Table 3	Examples of spectroscopic methods used for wine authentication	Pg 15
Table 4	The key areas in the wine supply chain determined by the GS1 traceability system	Pg 19

Chapter 2

Table 1	Examples of UV–visible spectroscopy in combination with chemometrics for wine authentication.	Pg 44
Table 2	Examples of IR spectroscopy in combination with chemometrics for wine authentication	Pg 45
Table 3	Examples of Raman spectroscopy in combination with chemometrics for wine authentication.	Pg 46
Table 4	Examples of fluorescence spectroscopy in combination with chemometrics for wine authentication	Pg 47
Table 5	Examples of NMR spectroscopy in combination with chemometrics for wine authentication.	Pg 48

Table 6	Summary of spectroscopic techniques applied to wine authentication [33,57]	Pg 49

Chapter 3

Table 1	Confusion matrix for cross-validated DA and XGBDA models from ICP-MS data for the different regions (abbreviations as defined in the Materials and methods section)	Pg 62
Table 2	Confusion matrix for SVMDA and XGBDA cross-validation results from EEM data for the different regions (Abbreviations as defined in the Materials and methods section)	Pg 64
Table 3	Least squares means and one-way ANOVA results of chemical measures within regions.	Pg 68
Table 4	Least squares means and one-way ANOVA results of selected elements for the different regions.	Pg 68
Table S1	XGBDA variable score plot for ICP-MS multi-element data giving a measure of a variable's importance for building the model	Pg 69
Table S2	Confusion matrices of different cross-validated models for the wine regions showing (a) DA of ICP-MS data, (b) XGBDA of ICP-MS data, and (c) XGBDA of EEM data.	Pg 69

Chapter 4

Table 1	Confusion matrix results for XGBDA analysis of multi-block data (absorbance + EEM) for wines (in duplicate) according to GI	Pg 77
Table 2	Confusion matrix results for XGBDA analysis of multi-block data (absorbance + EEM) for wines (in duplicate) according to variety.	Pg 78

List of Tables

Table S1	Summary of the different varieties and the regions (GIs) of the samples analysed in this study	Pg 84
Table S2	LS mean and one-way ANOVA results of basic chemical measures for the different varieties	Pg 84
Table S3	LS mean and one-way ANOVA results of chemical measures for the different regions	Pg 84
Table S4	Confusion matrix results for EEM data of XGBDA analysis according to region of origin.	Pg 85
Table S5	Confusion matrix results for EEM data of XGBDA analysis according to variety.	Pg 85
Table S6	Summary of statistics obtained from PLSR of HPLC reference values with A-TEEM measures	Pg 85
Table S7	LS mean and one-way ANOVA results of concentrations of phenolics for the varieties	Pg 86
Table S8	LS mean and one-way ANOVA results of concentrations of phenolics for the different regions	Pg 87
Table S9	Climatic data of region between 2018-2019 season	Pg 88

Chapter 5

Table 1	Descriptive statistics for basic chemical parameters of Shiraz wines (n = 186) from Barossa subregions for vintages 2018, 2019, and 2020	Pg 122
Table A1	Summary of the different vintages and region/subregions of the Shiraz wines	Pg 124
Table A2	Climatic data of the Barossa regions for year 2017-2020	Pg 124

Table A3	Confusion matrix results for XGBDA analysis of multi-block data (EEM and absorbance) for Shiraz wines modelled according to subregion and vintage	Pg 124
Table A4	Least squares means and one-way ANOVA results of basic chemical measures for the Shiraz wines.	Pg 125

Chapter 6

Table 1	Percentages of wine in blends of Shiraz with Cabernet Sauvignon or Merlot.	Pg 136
Table S1	Least squares means and one-way ANOVA results for oenological parameters and colour measurements (L*, lightness; a*, green–red; b*, blue–yellow; C*, chroma) of wine samples according to stage of winemaking.	Pg 146
Table S2	Least squares means and one-way ANOVA results for oenological parameters and colour measurements (L*, lightness; a*, green–red; b*, blue–yellow; C*, chroma) of wine samples according to origin for different varieties. Different letters within a column indicate statistically significant differences among the means according to Tukey HSD post hoc test (α = 0.05)	Pg 147

Appendix A

Publication

Modelling Cabernet Sauvignon wine sensory traits with spectrofluorometric data

Souza Gonzaga, L., Bastian, S. E. P., Capone, D. L., **Ranaweera, R. K. R.**, & Jeffery, D.W. (2021). Modelling Cabernet Sauvignon wine sensory traits with spectrofluorometric data. OENO One, 55 (4), 19-33.

OENO One
VINE AND WINE OPEN ACCESS JOURNAL

Received: 8 July 2021 | Accepted: 14 September 2021 | Published: 7 October 2021
DOI:10.20870/oeno-one.2021.55.4.4805

Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data

Lira Souza Gonzaga[1,2], Susan E. P. Bastian[1,2], Dimitra L. Capone[1,2], Ranaweera K. R. Ranaweera[1] and David W. Jeffery[1,2]*

[1] Department of Wine Science and Waite Research Institute, The University of Adelaide, PMB 1, Glen Osmond, South Australia 5064, Australia
[2]Australian Research Council Training Centre for Innovative Wine Production, The University of Adelaide, PMB 1, Glen Osmond, South Australia 5064, Australia

**corresponding author*: david.jeffery@adelaide.edu.au

Associate editor: Stamatina Kallithraka

ABSTRACT

Understanding how wine compositional traits can be related to sensory profiles is an important and ongoing challenge. Enhancing knowledge in this area could assist producers to select practices that deliver wines of the desired style and sensory specifications. This work reports the use of spectrofluorometry in conjunction with chemometrics for prediction, correlation, and classification based on sensory descriptors obtained using a rate-all-that-apply sensory assessment of Cabernet-Sauvignon wines (n = 26). Sensory results were first subjected to agglomerative hierarchical cluster analysis, which separated the wines into five clusters represented by different sensory profiles. The clusters were modelled in conjunction with excitation-emission matrix (EEM) data from fluorescence measurements using extreme gradient boosting discriminant analysis. This machine learning technique was able to classify the wines into the pre-defined sensory clusters with 100 % accuracy. Parallel factor analysis of the EEMs identified four main fluorophore components that were tentatively assigned as catechins, phenolic aldehydes, anthocyanins, and resveratrol (C1, C2, C3, and C4, respectively). Association of these four components with different sensory descriptors was possible through multiple factor analysis, with C1 relating to 'dark fruits' and 'savoury', C2 with 'barnyard', C3 with 'cooked vegetables' and 'vanilla/chocolate', and C4 with 'barnyard' and a lack of C1 descriptors. Partial least squares regression modelling was undertaken with EEM data and sensory results, with a model for perceived astringency being able to predict the panel scores with 68.1 % accuracy. These encouraging outcomes pave the way for further studies that relate sensory traits to fluorescence data and move research closer to the ultimate goal of predicting wine sensory expression from a small number of compositional factors.

KEYWORDS

Rate-all-that-apply, cluster analysis, excitation-emission matrix, partial least squares regression, machine learning, chemometrics

Supplementary data can be downloaded through: https://oeno-one.eu/article/view/4805

Appendix A | Modelling sensory traits with spectrofluorometry

Lira Souza Gonzaga *et al.*

INTRODUCTION

Wine is a luxury product with a highly complex composition that can be affected by the environment in which the grapes are grown as well as techniques applied in the vineyard and winery. The intrinsic complexity of wine has necessitated the development of various techniques to obtain an in-depth understanding of grape and wine metabolites and control points during production that can shape the final product. Relating compositional and technological factors with the sensory expression of a wine, which is a determining factor for the overall consumer experience, remains an ongoing focus of research. Being able to link chemical and sensory information with the practices and techniques that wine endures during production would ultimately equip practitioners with the ability to make more precise decisions for producing targeted wine styles.

Multiple methodologies are available for sensory profiling of wine, but their suitability will depend upon the requirements of the study. Rate-all-that-apply (RATA) is a quantitative sensory methodology that is rapid and effective for wine sensory characterisation (Danner *et al.,* 2018), as shown by its successful use in different studies (Franco-Luesma *et al.,* 2016; Mezei *et al.,* 2021; Nguyen *et al.,* 2020). Similarly to sensory profiling, a range of analytical approaches are available to define wine chemical composition that underpins sensory traits. A common approach has therefore been to combine sensory data with a number of chemical analysis techniques to predict and classify wine sensory characters (Niimi *et al.,* 2018), explore distinctiveness (Geffroy *et al.,* 2016), comprehend the impact of storage and packaging conditions (Hopfer *et al.,* 2013), and understand quality drivers (Gambetta *et al.,* 2016; Hopfer *et al.,* 2015). Many studies rely on analytical methodologies that are time-consuming, expensive, and relatively intricate (e.g., HPLC or GC with mass spectrometry), requiring personnel with specialised skills. There is room, however, for more accessible approaches (usually spectroscopy-based) that can provide chemical information more simply and rapidly. As reviewed by Ranaweera *et al.* (2021a), there are various spectroscopic approaches and each differs in terms of compounds measured, sensitivity, and advantages/disadvantages, among other aspects. The choice of methodology should therefore be defined according to the needs and objectives of the study.

As a spectroscopic technique, spectrofluorometry has often been applied to the analysis of food products because of its time- and cost-effective nature, and its high selectivity and sensitivity (Ranaweera *et al.,* 2021a). It can provide a unique three-dimensional excitation and emission matrix (EEM) that acts as a molecular fingerprint of a sample (Coelho *et al.,* 2015; Ranaweera *et al.,* 2021b). This technique can be a useful tool to authenticate, distinguish and classify different food products through a qualitative investigation of specific fluorescent substances (e.g., phenolic compounds, vitamins, and aromatic amino acids) present at different concentrations depending on the product (Karoui and Blecker, 2011). This methodology is also highly applicable to wine, which contains a myriad of fluorophores. Spectrofluorometry has been applied to wine for authentication and discrimination of samples based on variety, origin, or vintage (Ranaweera *et al.,* 2021b; Ranaweera *et al.,* 2021c; Sádecká and Jakubíková, 2020; Suciu *et al.,* 2019), to analyse oxidative changes and sulfur dioxide addition (Coelho *et al.,* 2015), and to quantitatively assess polyphenol content (Cabrera-Bañegil *et al.,* 2017).

In the quest for a rapid technique that could link wine composition and sensory properties, this study aimed to explore 1) the association between sensory descriptors obtained by RATA and the fluorescence EEM data recorded for Cabernet-Sauvignon wines from the Coonawarra Geographical Indication (GI), and 2) the dominant sensory traits of such regional wines. Specifically, the study tested the applicability of using EEMs with machine learning modelling for sample classification based on sensory profiles, investigated the relationship between the main fluorophores identified by parallel factor analysis (PARAFAC) and sensory descriptors using multiple factor analysis (MFA), and assessed partial least squares (PLS) regression models to predict sensory attributes.

MATERIALS AND METHODS

1. Sample selection

Unreleased vintage 2020 Cabernet-Sauvignon wines were sought from commercial producers using fruit from the Coonawarra GI of South Australia. Most of the wines were monovarietal and had only undergone alcoholic and malolactic fermentation and racking, with minimal oak contact (≤5 months) and limited maturation time.

In total, 26 Cabernet-Sauvignon wine samples (6 × 750 mL bottles of each wine) were obtained from 8 wineries/vineyards within the GI (Supplementary data, Table S1).

2. Sensory evaluation

Prior to formal evaluation, the wines were tasted by experts as defined by Parr *et al.* (2002) consisting of academics and postgraduate oenology students (n = 6), who evaluated aroma, flavour, taste, and mouthfeel with a free text assessment followed by a discussion of the wines. This informal tasting was used to evaluate whether the sample set was appropriate for a naïve panel to assess (considering that they were not commercially-released wines), to ensure that the samples could be differentiated, and to decide on the sensory attributes that should be included in the formal RATA evaluations.

Naïve wine consumers (n = 60; 27 females and 33 males from 18 to 77 years of age) were recruited based on being 18 years of age or older and having consumed red wine at least once a month. Evaluations were conducted in a purpose-built sensory laboratory at the University of Adelaide's Waite Campus, in individual booths equipped with a computer, under white fluorescent lighting, and at room temperature (22–23 °C). Samples (20 mL) were served at room temperature in clear stemmed ISO wine glasses coded with a random four-digit number and covered by a petri dish.

Due to the number of samples and to avoid palate fatigue, assessments were divided into three sessions: 9 samples in the first, 9 samples in the second, and 8 samples in the last session. The samples were randomly presented monadically for each subject within a session and the same panel was used for all three sessions. RATA methodology was used to characterise samples by rating the intensity only of the attributes that applied from a list of 53 comprising aroma, flavour, taste, and mouthfeel descriptors (Supplementary data, Table S2) on a 7-point scale (from "extremely low" to "extremely high"). Between samples, the panellists were forced to have a 1-min break and could cleanse their palate with deionised water and unsalted crackers. A 5-min break was enforced at the mid-point of the tasting (between samples 4 and 5). Data were collected with RedJade software (2016, Redwood City, USA). Informed consent was obtained from panellists and this study was approved by the Human Research Ethics Committee of the University of Adelaide (approval number: H-2019-031).

3. Chemicals

HPLC grade absolute ethanol and analytical grade 37 % hydrochloric acid (HCl) were purchased from Chem-Supply (Port Adelaide, SA, Australia). High purity water was obtained from a Milli-Q purification system (Millipore, North Ryde, NSW, Australia).

4. Spectrofluorometric analysis

After sensory analysis, the remainder of each wine was subsampled into a 4 mL centrifuge tube that was completely filled and stored in a refrigerator at 4 °C until measurements were performed. After warming to room temperature, samples were centrifuged at 9300 × *g* for 10 min and diluted with 50 % aqueous ethanol that had been adjusted with HCl to pH 2 and vacuum filtered (0.45 µm PTFE membrane). The samples were diluted 150-fold (Ranaweera *et al.,* 2021c), and analysed in a Hellma type 1FL (1 cm path length) Macro Fluorescence cuvette (Sigma-Aldrich, Castle Hill, NSW, Australia). Samples were prepared in duplicate and two measurements of each sample were undertaken with a Horiba Scientific Aqualog® spectrophotometer (version 4.2, Quark Photonics, Adelaide, SA, Australia). The excitation wavelength ranged from 240 to 700 nm with an increment of 5 nm under medium gain and 0.2 s integration time and the emission wavelength ranged from 242 to 824 nm with an increment of 4.66 nm. Data acquisition was controlled with Origin software (version 8.6, OriginLab® Corporation, Massachusetts, USA) and EEMs were normalised using water Raman scattering units and corrected for the inner filter effects, solvent background, dark detector signals, and Rayleigh masking (Gilmore *et al.,* 2017).

5. Basic analytical measurements of pH, TA, ethanol, and SO_2

Sample pH and titratable acidity (TA) were obtained with a T50 auto-titrator (Mettler Toledo, Melbourne, VIC, Australia). Ethanol was measured in triplicate by HPLC analysis (Li *et al.,* 2017) of undiluted samples that were centrifuged at 9300 × *g* for 10 min. Separation was performed with an Aminex HPX-87H column (300 mm × 7.8 mm, BioRad, Hercules, California, USA) thermostatted at 60 °C using 2.5 mM H_2SO_4 as mobile phase with a flow rate of 0.5 $mLmin^{-1}$. Peaks were detected with a refractive index detector (RID-10A, Shimadzu, Kyoto, Japan) and quantified by comparison with standards prepared in model wine using ChemStation for

LC 3D Systems software (Agilent Technologies, Santa Clara, CA, USA). Free and total SO_2 concentrations were determined in duplicate using the method described by Iland *et al.* (2004).

6. Statistical analysis

The raw sensory data were firstly analysed through two-way analysis of variance (ANOVA) with panellists as a random factor and samples as a fixed factor to identify significantly different attributes between the samples. Attributes that presented a p-value ≤ 0.1 were selected for agglomerative hierarchical cluster (AHC) analysis of all samples with an automatic entropy truncation and Euclidean distance using Ward's method or unweighted pair-group average (UPGMA). With a superior cophenetic correlation (0.676 for UPGMA *versus* 0.511 for Ward's method), UPGMA was chosen and truncation configured with a minimum of five classes. Correlation principal component analysis (PCA) was performed to identify sensory profiles that arose for different clusters based on the AHC analysis.

EEM data were unfolded using unfold multiway (mode 1) in Solo software (version 8.7.1, Eigenvector Research, Inc., Manson, WA, USA). For classification according to the clusters defined by AHC analysis, extreme gradient boosting discriminant analysis (XGBDA) was conducted (Ranaweera *et al.,* 2021c) using pre-processing with mean centring, PLS compression to yield a maximum of 25 latent variables (LVs), and decluttering with generalised least squares weighting at 0.2 for calibration and cross-validation (k = 10, Venetian blinds procedure). Confusion matrix score probabilities were used to assess the model effectiveness. PARAFAC was performed with a non-negativity constraint in all modes imposed and the model was validated by split-half analysis (Murphy *et al.,* 2013).

Loadings for the components determined by PARAFAC were analysed in conjunction with the sensory data (significantly different attributes, α = 0.1) through MFA. Separately, a calibration model was created with PLS1 regression of sensory scores for perceived wine astringency and the EEM data to predict astringency ratings. The model was optimised through assessment of LVs, root mean square error of calibration (RMSEC), root mean square error of cross-validation (RMSECV, Venetian blinds with 10 splits), and root mean square error of prediction (RMSEP).

ANOVA, PCA, AHC, and MFA were performed with XLSTAT (version 2019.4.1, Addinsoft, New York, USA). XGBDA, PARAFAC, and PLS regression analysis were conducted with Solo software (version 8.7.1).

RESULTS AND DISCUSSION

Unreleased Cabernet-Sauvignon wines sought for the study went through minimal post-fermentation processes (e.g., fining, maturation, blending) and were bottled at early stages of production so that the impact of the Coonawarra GI could be assessed with minimal influence of downstream winemaking operations. Basic analytical measurements were within the normal range for red wines at such a stage of production. The total and free SO_2 content ranged from 0.4 to 70.8 mgL^{-1} and 0.4 to 33.4 mgL^{-1}, respectively, TA ranged from 5.6 to 7.5 gL^{-1}, pH values ranged from 3.40 to 3.87, and ethanol concentration ranged from 12.9 % to 15.3 % (Supplementary data, Table S1).

1. RATA sensory profiling and clustering of wines

Of the 53 sensory attributes rated by panellists using RATA methodology, 20 were significantly different (α = 0.1) according to ANOVA and comprised 8 aromas, 8 flavours, 3 tastes, and 1 mouthfeel attribute (Supplementary data, Table S3). The means of the 20 descriptors were analysed through a correlation PCA (Figure 1) following the AHC analysis (Supplementary data, Figure S1). The first factor (F1) in Figure 1A accounted for 30.6 % of the data variance and the second factor (F2) explained a further 19.6 %. Cluster 1 (shown in red, 7 wines) appeared on the right side of F1 and spread across both segments of F2, with 5 samples in the upper half and 2 in the lower half. Cluster 2 (green, 14 samples) mostly presented near the origin, with 11 samples on the left and 3 samples on the right of F1, and a more or less even spread across F2. Cluster 3 (cyan, 2 samples) was found on the left side of F1 and upper half of F2, and Cluster 4 (pink, 1 sample) was separated from the rest in the bottom right portion of the plot. Squared cosine values for samples in Cluster 5 (data not shown) indicated a higher representation on F3, in the lower half as seen in Figure 1B.

In terms of the sensory descriptors, 'barnyard' flavour and aroma, and bitterness and astringency were plotted on the right side of F1 and lower part of F2; 'minty', 'cooked vegetables', 'dark fruits', 'tobacco', and 'earthy' aromas and flavours, 'oaky' and 'savoury' aromas, and acidity were plotted

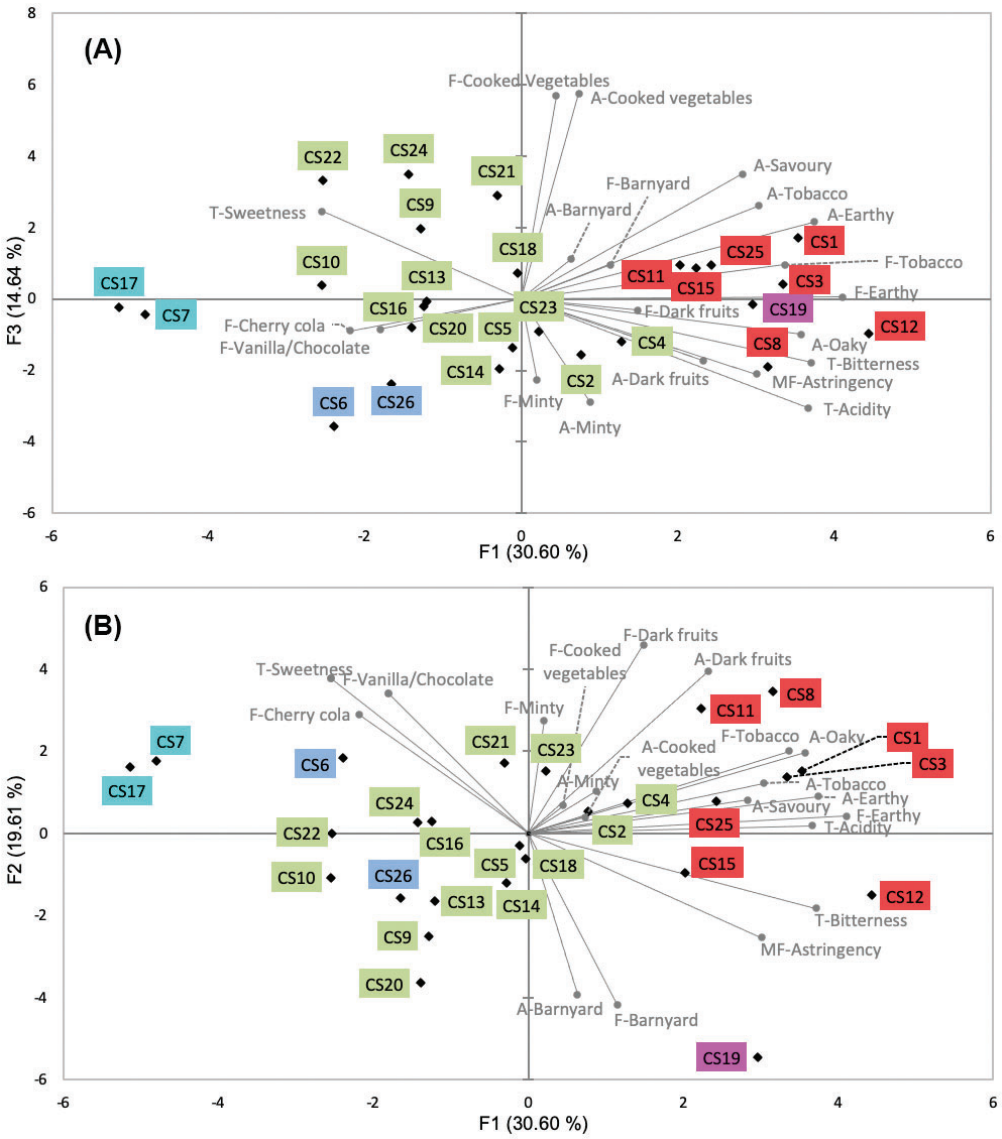


FIGURE 1. Principal component analysis biplots of Cabernet-Sauvignon wines (n = 26) using significantly different (α = 0.1) RATA attributes, showing (A) F1 *versus* F2 and (B) F1 *versus* F3.

Colour coding represents the clusters resulting from the agglomerative hierarchical cluster analysis (Supplementary data, Figure S1), with samples in the same cluster bearing the same colour. Cluster 1, red; Cluster 2, green; Cluster 3, cyan; Cluster 4, pink; Cluster 5, blue. A-, aroma; F-, flavour; MF-, mouthfeel; T-, taste.

on the right side of F1 and upper half of F2; and 'vanilla/chocolate' and 'cherry cola' flavours, and sweetness were plotted on the left side of F1 and upper half of F2 (Figure 1A). The aroma and flavour of 'cooked vegetables' were better represented in the upper half of F3 (Figure 1B).

The clusters defined by AHC analysis (Supplementary data, Figure S1) could be explained through different sensory profiles as shown in Figure 1. Cluster 1 was characterised by savoury characters including 'earthy' and 'tobacco', along with 'oaky' and 'dark fruits' aromas, and higher acidity, whereas Cluster 2 on the opposite side was generally characterised by a lack of those characters. Considering that these were young wines, the results might indicate the presence of some oak contact during fermentation for most samples in Cluster 1 as opposed to no oak contact for samples in Cluster 2 (Crump *et al.,* 2015).

Cluster 3 was associated with higher sweetness and 'cherry cola' flavour and low bitterness and astringency. Cluster 4 was characterised by 'barnyard' aroma and flavour, relatively low 'vanilla/chocolate' and 'cherry cola' flavours, a higher bitter taste and astringent mouthfeel, and a lack of sweetness. Cluster 5 was especially related to 'cherry cola' and 'vanilla/chocolate' flavours (Figure 1B), as opposed to the savoury profile found for Cluster 1 (Figure 1A). Sensory profiles have similarly been used in the past for regional classification of Australian Cabernet-Sauvignon wines (Souza Gonzaga *et al.,* 2019; Souza Gonzaga *et al.,* 2020) and Australian Shiraz and Chardonnay wines (Kustos *et al.,* 2020). Those studies with commercial wines reported that some distinctive sensory traits can be more important and more associated with a specific wine-producing region, with the current work on unreleased wines also indicating the existence of perceived differences within a GI according to Figure 1.

The main differences reported previously for Cabernet-Sauvignon wines were the duality between 'green' and 'fruity' related characters and between 'oak' related traits and 'eucalyptus' or 'minty' attributes (Heymann and Noble, 1987; Souza Gonzaga *et al.,* 2020). In the present study, the contrast was between 'barnyard', astringency and bitterness attributes, and 'cherry cola', 'vanilla/chocolate', and sweetness. Oak-related and savoury attributes and the 'minty' trait were found in the same quadrant, not in direct contrast, and the same was evident for fruity and vegetal characters (Figure 1A). Considering the samples were dominated by or exclusively produced from Cabernet-Sauvignon (Supplementary data, Table S1) and were all from the same GI, albeit from different vineyards and wineries, the disparity in the sensory profiles of the present work might be associated with differences in the winemaking processes, as seen previously by Kustos *et al.* (2020) with Australian Chardonnay and Shiraz wines. Additionally, the wines in the present study had a minimal influence of oak (i.e., less than 5 months) or other maturation treatments compared to commercially released red wines, which might have allowed sensory traits that could be attributed to aspects of terroir (e.g., soil, topography, and vineyard management practices) to be more perceivable, such as the 'minty' and fruity attributes.

Some samples in Cluster 2 indicated that 'minty' flavour was an important characteristic, although in general not much difference was seen between the samples (Figure 1A). A 'minty' character has been reported previously for Coonawarra Cabernet-Sauvignon wines, which might indicate this as a dominant trait for the Coonawarra region (Robinson *et al.,* 2011; Souza Gonzaga *et al.,* 2019; Souza Gonzaga *et al.,* 2020). Characters described as 'minty' and 'eucalyptus' in Cabernet-Sauvignon wines have been associated with the presence of eucalyptol (i.e., 1,8-cineole) and hydroxycitronellol, and although 'eucalyptus' aroma and flavour were not statistically significant ($\alpha = 0.1$) in the present work (Supplementary data, Table S3), studies have shown that they might be interchangeable and indistinguishable by a sensory panel (Capone *et al.,* 2012; Robinson *et al.,* 2011; Souza Gonzaga *et al.,* 2020). The current study did not explore the presence of volatile compounds so the link between 'minty' and 'eucalyptus' from both sensory and chemical viewpoints is open for further examination. Among the possibilities, the occurrence of 1,8-cineole in wine has been related to the presence of *Eucalyptus* trees within the vineyard environment (Capone *et al.,* 2012), whereas some studies report the presence of 'minty' traits associated with an aged profile of Bordeaux red wines specifically under the influence of the proportion of Cabernet-Sauvignon in the blend (Picard *et al.,* 2015; Picard *et al.,* 2016b). Mint aroma in that case has been associated with the presence of piperitone (Picard *et al.,* 2016a). Considering that the present study examined young Cabernet-Sauvignon wines, it seemed unlikely that piperitone or other limonene-derived compounds (Picard *et al.,* 2017) were responsible for the presence of the 'minty' attribute, although further investigation is required to clarify the role of various monoterpenoids in the perception of mint-related characters.

2. Classification of sensory clusters based on spectrofluorometric analysis

To examine whether sensory information could be classified using spectrofluorometric data, the results from AHC (Supplementary data, Figure S1) were modelled in conjunction with the EEMs of the wine samples through machine learning with the XGBDA algorithm. Various algorithms and machine learning tools exist for wine classification based on EEM data, such as soft independent modelling of class analogy and support vector machine, but XGBDA performs well when analysing a complex heterogeneous matrix with uneven class distribution (Babajide Mustapha and Saeed, 2016). The analysis was undertaken

after PLS compression, used to improve the stability of the model by making it less disposed to overfitting. The class CV prediction demonstrated in Figure 2 shows each cluster (denoted using different symbols and colours) that was predefined by AHC. The model attempted to predict the class (cluster) to which each sample belonged, based on the relationship of the sensory profiles and EEM data. Figure 2 and the confusion matrix obtained from cross-validation (data not shown) highlighted that all clusters were 100 % correctly classified with a discrete segregation between the classes in the cross-validated model. This result indicated that the underlying composition of the wines encompassed in the fluorescence fingerprints might be driving the sensory differences of the clusters determined from RATA evaluation.

Classification methods using fluorescence spectroscopy have been previously applied for wine varietal, vintage and origin authentication (Ranaweera *et al.,* 2021b; Ranaweera *et al.,* 2021c; Sádecká and Jakubíková, 2020; Suciu *et al.,* 2019), which tends to yield similar or even better performance compared to other spectroscopic methods like UV-vis, near-infrared, mid-infrared, synchronous fluorescence, or Raman (Mandrile *et al.,* 2016; Riovanto *et al.,* 2011; Tan *et al.,* 2016). Ultimately, studies involving spectrofluorometry and chemometrics have demonstrated the approach as a valid tool for authenticating wine, and along with the present work, highlight the extent to which this type of data can be used to understand important traits related to wine chemical and sensory properties.

3. Using PARAFAC to identify main fluorophoric compounds

Attempting to shed light on the relationship between fluorescence data and sensory properties, PARAFAC was performed on the EEM data to identify the main fluorophores present in the samples. The percentage of core consistency of the data can be applied in combination with split-half analysis to assess the model suitability, especially with high complexity matrices such as wine (Airado-Rodríguez *et al.,* 2011; Murphy *et al.,* 2013). The split-half analysis compares the similarity between each half of

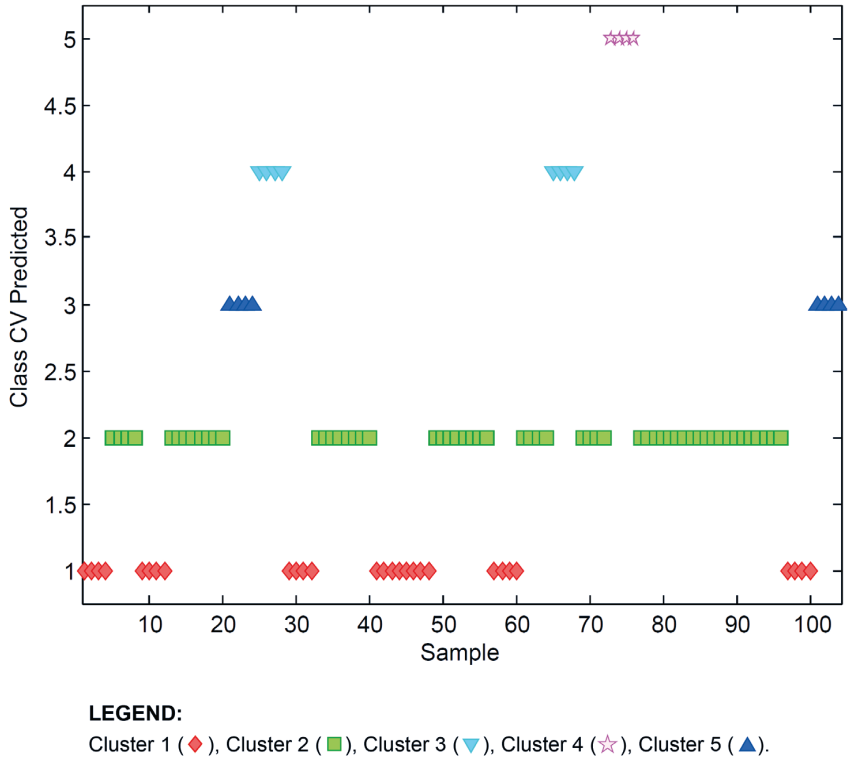


FIGURE 2. Class CV predicted for classification of RATA clusters arising from AHC based on XGBDA modelling for the set of Cabernet-Sauvignon wines (n = 26).

the data set, and like with core consistency, a higher percentage is desirable when deciding on the number of components for the model (Murphy *et al.,* 2013). Using all samples in the first PARAFAC model generated a core consistency of less than 0 % and a split-half result of less than 19 %. Investigating further, analysis of residuals of the samples showed that three (CS2, CS7 and CS26) of the 26 wines were outliers and presented equally high residuals for the four determinations (i.e., duplicate readings of duplicate samples) compared to the other samples. Based on the available data, no possible reason was identified that could explain the three samples as outliers. Although sample CS7 was the only sample produced with 100 % uninoculated alcoholic and malolactic fermentation, which might indicate a possible factor, that was not the case for the other two outlier samples. Nonetheless, PARAFAC modelling was performed again without the outlier samples, this time yielding a core consistency of 61 % and split half analysis of 93.7 % for the four main fluorescent components (Figure 3).

From PARAFAC it was possible to identify the maximum intensities (λ_{ex} and λ_{em}) for the four components as demonstrated in Figure 3, and therefore to tentatively assign chemical compound classes that are naturally present in wine (Airado-Rodríguez *et al.,* 2011; Airado-Rodríguez *et al.,* 2009). Such spectral data can typically be related to fluorophoric compounds such as vitamins (Christensen *et al.,* 2006) and especially phenolic compounds (Schueuermann *et al.,* 2018). For PARAFAC component 1, maximum intensities of λ_{ex} = 275 nm and λ_{em} = 310 nm were tentatively

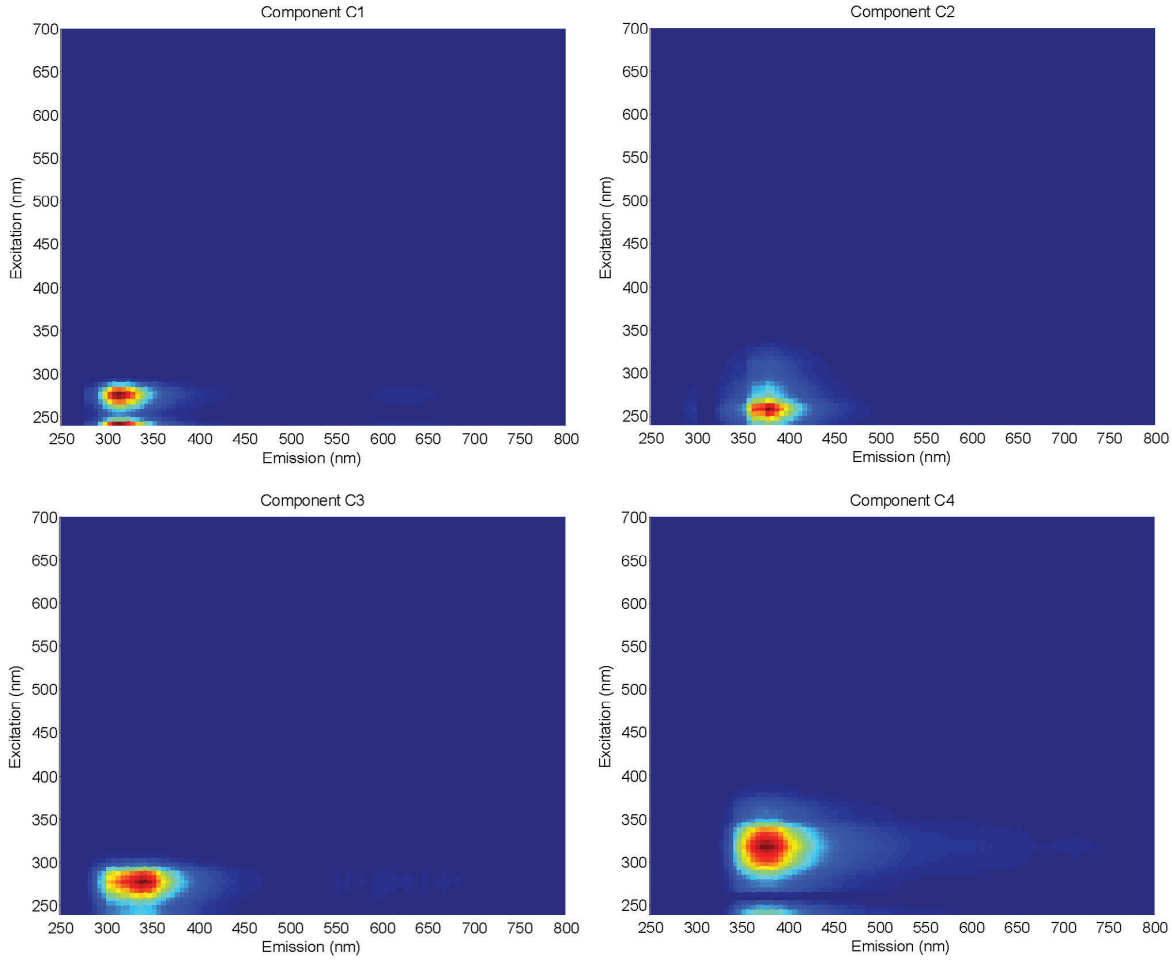


FIGURE 3. Contour plots for excitation and emission wavelengths identified from the PARAFAC model, indicating the four main fluorescent components (i.e., C1, C2, C3, C4) present in the sample set.

identified as compounds associated with catechin (including tannin). Component 2 peak intensities were λ_{ex} = 255 nm and λ_{em} = 375 nm and can be proposed to result from phenolic aldehyde related compounds. Component 3 peak intensities were λ_{ex} = 270 nm and λ_{em} = 335 nm and were considered to be associated with anthocyanins. Finally, component 4 peak intensities were λ_{ex} = 315 nm and λ_{em} = 375 nm and tentatively assigned to stilbenoids such as *trans*-resveratrol.

Ranaweera *et al.* (2021c) and Airado-Rodríguez *et al.* (2009) proposed similar assignments for PARAFAC model components in red wine, which are reasonable considering the main compounds (i.e., catechins, anthocyanins, and other phenolics) expected to be abundant in red wine. It is noteworthy that compound classes assigned from the PARAFAC modelling (i.e., phenolics) were not necessarily driving the sensory characters themselves, but could act as indirect markers that indicated compositional aspects of the wines that were not essentially measured by fluorescence. For example, different gene copies responsible for the biosynthesis of important wine compounds such as anthocyanins in grape berry can belong to multicopy families, having an expression profile coinciding with other specific flavonoids that may impact wine sensory profile by correlation rather than causation (Kuhn *et al.,* 2013). In contrast, there could be a direct relationship with compounds associated with aspects such as the taste and mouthfeel of the wine, as explained in more detail in the next section.

4. Relation between PARAFAC components and RATA results according to MFA

Considering the compound classes tentatively identified by PARAFAC modelling of EEM data can impact wine sensory profile (either directly or by implying an indirect correlation), the relative loadings of the four classes were analysed in conjunction with RATA results through MFA. Means of the significantly different (α = 0.1) descriptors and means of the four compound class loadings from 23 wines (excluding CS2, CS7 and CS26) were used for the analysis (Figure 4). MFA yielded an RV coefficient of 0.232 between both sets of data, an RV coefficient of 0.751 between PARAFAC data and the MFA model, and an RV coefficient of 0.816 between the RATA data and the MFA model.

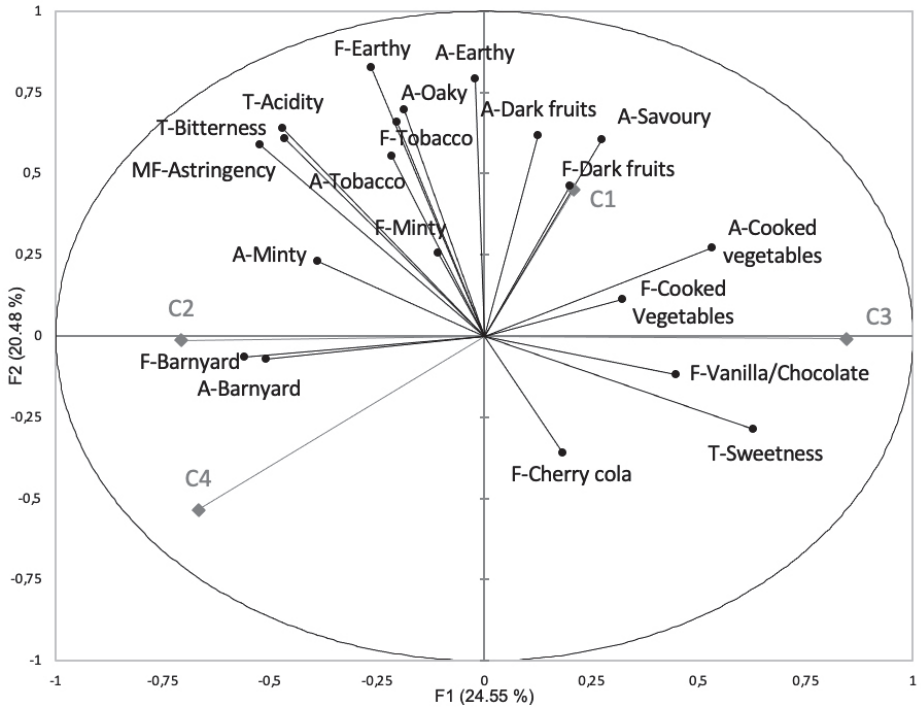


FIGURE 4. Multiple factor analysis biplot of the four components from PARAFAC (in grey, ◆) using significantly different (α = 0.1) descriptors from RATA evaluation (in black, ●) for 23 Cabernet-Sauvignon wine samples (excluding CS2, CS7 and CS26).

The MFA biplot explained 45 % of the variance in the data, with 24.6 % represented by F1 and 20.5 % by F2. PARAFAC C1 was plotted on the right side of F1 and the upper portion of F2, C2 and C3 were explained entirely along F1, with C3 on the right side and C2 on the left side, and C4 was plotted on the left side of F1 and lower part of F2, more or less opposite to C1 (Figure 4).

Catechin monomers associated with C1 are usually extracted from grape skin and seed and can increase the bitter taste of wine (Fischer and Noble, 1994) whereas polymers of catechin (e.g., tannins), extracted from the same sources, are related with astringency (Waterhouse *et al.,* 2016a). Figure 4 shows C1 was associated with 'dark fruits' and 'cooked vegetables' aromas and flavours and 'savoury' aroma, which is likely to be an indirect relationship as mentioned in the previous section. Analysing the RV coefficients, the correlation between bitterness and C1 was not significant ($p = 0.313$), thus indicating that there might not be an association. In contrast, the correlation between astringency and C1 was significant ($p = 0.006$) and had an RV coefficient of 0.315, demonstrating a moderate association. This implied that polymers had a greater influence on the expression of C1 than monomers, which would be reasonable given their relative concentrations in red wine.

Phenolic aldehydes assigned to C2 can be influenced by the origin of wood (usually oak) incorporated either during fermentation or maturation and can vary in concentration depending on ageing time — such compounds can be responsible for some oak-related aroma traits (e.g., vanillin) in wine (del Alamo Sanza *et al.,* 2004). Other oak compounds (e.g., volatile phenols, hydrolysable tannins) that may influence sensory traits would undoubtedly be extracted as well. C2 was related to 'barnyard' aroma and flavour and 'minty' aroma. Anthocyanins assigned to C3 are pigments present in red grape skins that are important to the colour of red wine (He *et al.,* 2012). Anthocyanins might also be responsible for an increase in the 'fullness' of a wine (Vidal *et al.,* 2004), as well as perceived astringency and bitterness (Ferrero-del-Teso *et al.,* 2020; Paissoni *et al.,* 2018). Additionally, as explained in the section dealing with PARAFAC, genes involved in the biosynthesis of anthocyanins in grapes are expressed through pathways that coincide with the biosynthesis of other flavonoids and volatile compounds (Czemmel *et al.,* 2012; Kuhn *et al.,* 2013). This could explain why anthocyanins could act as markers for compounds that impart aroma or flavour (Ristic *et al.,* 2010) but lack a fluorophore themselves. From the MFA, C3 was linked to 'cooked vegetables' aroma and flavour, 'vanilla/chocolate' flavour, and sweetness. Lastly, stilbenoids assigned to C4 are compounds that can be found in grape berry skins and are extracted into wine during fermentation (Waterhouse *et al.,* 2016b). Stilbenoids, especially *trans*-resveratrol, are responsible for the antioxidant characteristics of red wine and its association with the prevention of age-related diseases in consumers (Pawlus *et al.,* 2012). According to Gaudette and Pickering (2011), *trans*-resveratrol seems to have minimal impact on the sensory qualities of wine (when spiked at less than 200 mgL^{-1}). Figure 4 shows that C4 was associated with 'barnyard' aroma and flavour, which is likely to be another example of an indirect relationship between the fluorophoric component and the sensory data.

It is worth noting that the associations between sensory traits and tentative compound types found through PARAFAC do not allow for strict conclusions. It is possible, considering the complexity of what is being modelled, that some relationships may arise due to chance, and more in-depth research is necessary to better understand and explain the proposed relationships.

5. Regression model for astringency prediction

Considering that most of the compounds detected by spectrofluorometric analysis can directly affect basic mouthfeel and taste attributes in wine, PLS regression was performed with the two mouthfeel and three taste attributes described by the sensory evaluation of the 26 wines. Astringency was the only attribute that could be well modelled from the EEM data without overfitting, based on the model parameters. An optimal model was generated with eight LVs, giving RMSEC = 0.085, RMSECV = 0.132, RMSEP = 0.222, R^2 calibration = 0.936, R^2 cross-validation = 0.848, and R^2 prediction = 0.681. The model was thus able to explain 84.8 % of the variance in the samples and able to predict the results with 68.1 % accuracy (Figure 5). Furthermore, the low value for RMSECV indicated that the error associated with the prediction of astringency was around 2 % in relation to the sensory scale used (7-point), demonstrating that the model appeared to be suitable. This outcome showed that spectrofluorometric data had reasonable capabilities for predicting a perceived mouthfeel attribute rating for this data set, which was

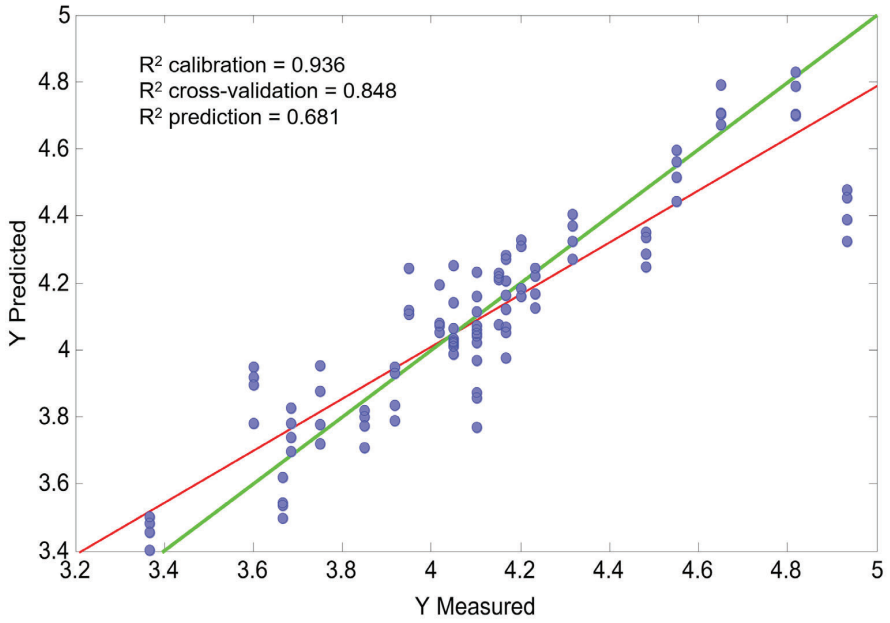


FIGURE 5. Correlation between the predicted and measured ratings for perceived astringency according to partial least squares regression modelling for Cabernet-Sauvignon wines (n = 26).
The green line shows the 1:1 correlation and the red line is the model fit.

encouraging given the simplicity of the approach and the complexity of what was being modelled.

The chemical composition of Cabernet-Sauvignon wines has also previously been used for sensory profile prediction, with regression models described by Niimi *et al.* (2018) explaining between 44.2 % and 69.1 % of the variance in the sample set, and 56.5 % for astringent mouthfeel. In that work, the model for predicting perceived astringency score involved anthocyanin concentration and colour measures, both of which can be determined using the A-TEEM approach and used in combination with a multi-block analysis (Ranaweera *et al.,* 2021c) to add information beyond that encompassed in the EEM data alone. Notably, the present study is the first known attempt to correlate and predict wine sensory profiles from EEM readings, and although the outcomes are positive, further work with additional samples will be necessary to improve and extend the modelling. Furthermore, different spectroscopic methods have been validated before for determining phenolic compound concentrations in a way that is less time consuming and more cost-effective than other options, and such approaches could become a valuable tool for assisting winemakers in monitoring and controlling phenolic composition (Cozzolino *et al.,* 2008; Cozzolino *et al.,* 2004; Dambergs *et al.,* 2012; Janik *et al.,* 2007; Ranaweera *et al.,* 2021c). Fluorescence spectroscopy in particular can quantify compounds that are present in the sample at a lower concentration than other spectroscopic methods (Gilmore and Chen, 2020), thus providing an attractive option for additional development in future.

CONCLUSIONS

This study aimed to explore the association between sensory traits and spectrofluorometric data of unreleased, commercially produced 2020 Coonawarra Cabernet-Sauvignon wines. It combined cluster analysis of sensory profiles obtained using RATA with fluorescence data by using a machine learning algorithm, and examined the prediction of sensory ratings from fluorophoric compounds via regression modelling. Thus, five distinctive clusters arose that could be well explained by the sensory results of the RATA evaluation. Cluster 1 wines were characterised by savoury-related characters, Cluster 2 by 'minty' traits and a lack of the savoury-related attributes, Cluster 3 by 'cherry cola' flavour and low bitterness and astringency, Cluster 4 by higher sweetness and 'barnyard' aroma and flavour, and Cluster 5 by 'vanilla/chocolate' flavour.

Additionally, the EEM data analysed through XGBDA were able to predict with 100 % accuracy the clusters that arose from the sensory profiling, demonstrating that there might be a good association between the EEMs and sensory ratings (whether direct or indirect). After excluding three outlier samples, PARAFAC analysis showed that four main fluorophores could be segregated to explain the data set, with compound classes tentatively associated with the intensity readings being catechins (C1), phenolic aldehydes (C2), anthocyanins (C3) and stilbenoids (C4). MFA was used to identify associations between the PARAFAC components and the sensory ratings, revealing that C1 was associated with 'dark fruits' and 'savoury' characters, C2 was associated with 'barnyard', C3 was related to 'cooked vegetables' and 'vanilla/chocolate', and C4 was related with 'barnyard' but more characterised by the lack of attributes associated with C1. However, the nature of any relationship between the proposed compound classes and perceived sensory attributes requires further study. PLS regression resulted in a suitable model that was able to predict perceived astringency score with 68.1 % accuracy, although no suitable model was found for the other sensory attributes. Overall, the correlation of sensory profiles with spectrofluorometric data was quite an optimistic feat, yet the results from this study were promising. This work may inspire further research that is designed to better understand the chemical drivers of sensory traits and the most influential factors throughout wine production using a rapid technique like spectrofluorometry, perhaps with the inclusion of a small selection of compositional variables.

Acknowledgements: We are especially grateful to Coonawarra Vignerons for their valuable input, support and organising the donation of wines. The authors thank ARC Training Centre for Innovative Wine Production and Department of Wine Science colleagues at the University of Adelaide for their encouragement and support during the trial and appreciate the time committed by panellists. In particular, the authors appreciate the technical assistance from Claire Armstrong. This research was conducted by the Australian Research Council Training Centre for Innovative Wine Production (www.ARCwinecentre.org.au; project number IC170100008), funded by the Australian Government with additional support from Wine Australia, Waite Research Institute and industry partners. Wine Australia also provided a supplementary scholarship to L.S.G (WA Ph1804) and to R.K.R.R (WA Ph1909). The University of Adelaide is a member of the Wine Innovation Cluster.

REFERENCES

Airado-Rodríguez, D., 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*(2), 257-264. https://doi.org/10.1016/j.jfca.2010.10.005

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*(5), 1711-1720. https://doi.org/10.1021/jf8033623

Babajide Mustapha, I., & Saeed, F. (2016). Bioactive molecule prediction using extreme gradient boosting. *Molecules, 21*(8), 983. https://doi.org/10.3390/molecules21080983

Cabrera-Bañegil, M., Hurtado-Sánchez, M. d. 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. https://doi.org/10.1016/j.foodchem.2016.09.152

Capone, D. L., Jeffery, D. W., & Sefton, M. A. (2012). Vineyard and fermentation studies to elucidate the origin of 1,8-cineole in Australian red wine. *Journal of Agriculture and Food Chemistry, 60*(9), 2281-2287. https://doi.org/10.1021/jf204499h

Christensen, J., Nørgaard, L., Bro, R., & Engelsen, S. B. (2006). Multivariate autofluorescence of intact food systems. *Chemical Reviews, 106*(6), 1979-1994. https://doi.org/10.1021/cr050019q

Coelho, C., Aron, A., Roullier-Gall, C., Gonsior, M., Schmitt-Kopplin, P., & Gougeon, R. D. (2015). Fluorescence fingerprinting of bottled white wines can reveal memories related to sulfur dioxide treatments of the must. *Analytical Chemistry, 87*(16), 8132-8137. https://doi.org/10.1021/acs.analchem.5b00388

Cozzolino, D., Cynkar, W. U., Dambergs, R. G., Mercurio, M. D., & Smith, P. A. (2008). Measurement of condensed tannins and dry matter in red grape homogenates using near infrared spectroscopy and partial least squares. *Journal of Agricultural and Food Chemistry, 56*(17), 7631-7636. https://doi.org/10.1021/jf801563z

Cozzolino, D., Kwiatkowski, M. J., Parker, M., Cynkar, W. U., Dambergs, R. G., Gishen, M., & Herderich, M. J. (2004). Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta, 513*(1), 73-80. https://doi.org/10.1016/j.aca.2003.08.066

Crump, A. M., Johnson, T. E., Wilkinson, K. L., & Bastian, S. E. P. (2015). Influence of oak maturation regimen on composition, sensory properties, quality, and consumer acceptability of Cabernet-Sauvignon

wines. *Journal of Agricultural and Food Chemistry, 63*(5), 1593-1600. https://doi.org/10.1021/jf5044025

Czemmel, S., Heppel, S. C., & Bogs, J. (2012). R2R3 MYB transcription factors: key regulators of the flavonoid biosynthetic pathway in grapevine. *Protoplasma, 249*(2), 109-118. https://doi.org/10.1007/s00709-012-0380-z

Dambergs, R. G., Mercurio, M. D., Kassara, S., Cozzolino, D., & Smith, P. A. (2012). Rapid measurement of methyl cellulose precipitable tannins using ultraviolet spectroscopy with chemometrics: application to red wine and inter-laboratory calibration transfer. *Applied Spectroscopy, 66*(6), 656-664. https://doi.org/10.1366/11-06516

Danner, L., Crump, A. M., Croker, A., Gambetta, J. M., Johnson, T. E., & Bastian, S. E. P. (2018). Comparison of Rate-All-That-Apply and descriptive analysis for the sensory profiling of wine. *American Journal of Enology and Viticulture, 69*(1), 12-21. https://doi.org/10.5344/ajev.2017.17052

del Alamo Sanza, M., Nevares Domínguez, I., Cárcel Cárcel, L. M., & Navas Gracia, L. (2004). Analysis for low molecular weight phenolic compounds in a red wine aged in oak chips. *Analytica Chimica Acta, 513*(1), 229-237. https://doi.org/10.1016/j.aca.2003.11.041

Ferrero-del-Teso, S., Suárez, A., Jeffery, D. W., Ferreira, V., Fernández-Zurbano, P., & Sáenz-Navajas, M.-P. (2020). Sensory variability associated with anthocyanic and tannic fractions isolated from red wines. *Food Research International, 136*, 109340. https://doi.org/10.1016/j.foodres.2020.109340

Fischer, U., & Noble, A. C. (1994). The effect of ethanol, catechin concentration, and pH on sourness and bitterness of wine. *American Journal of Enology and Viticulture, 45*(1), 6. http://www.ajevonline.org/content/45/1/6.abstract

Franco-Luesma, E., Sáenz-Navajas, M.-P., Valentin, D., Ballester, J., Rodrigues, H., & Ferreira, V. (2016). Study of the effect of H_2S, MeSH and DMS on the sensory profile of wine model solutions by Rate-All-That-Apply (RATA). *Food Research International, 87*, 152-160. https://doi.org/10.1016/j.foodres.2016.07.004

Gambetta, J. M., Cozzolino, D., Bastian, S. E. P., & Jeffery, D. W. (2016). Towards the creation of a wine quality prediction index: Correlation of Chardonnay juice and wine compositions from different regions and quality levels. *Food Analytical Methods, 9*(10), 2842-2855. https://doi.org/10.1007/s12161-016-0467-9

Gaudette, N. J., & Pickering, G. J. (2011). Sensory and chemical characteristics of *trans*-resveratrol-fortified wine. *Australian Journal of Grape and Wine Research, 17*(2), 249-257. https://doi.org/10.1111/j.1755-0238.2011.00144.x

Geffroy, O., Buissière, C., Lempereur, V., & Chatelet, B. (2016). A sensory, chemical and consumer study of the peppery typicality of French Gamay wines from cool-climate vineyards. *Journal International de la Vigne et du Vin, 50*(1), 35-47. https://hal.archives-ouvertes.fr/hal-01308386

Gilmore, A. M., Akaji S., & Csatorday, K. (2017). Spectroscopic analysis of red wines with A-TEEM molecular fingerprinting. *Readout, 49*, 41-48.

Gilmore, A. M., & Chen, L. (2020). *Water soluble fraction (WSF) contaminant detection using machine-learning absorbance-transmission excitation emission matrix (A-TEEM) spectroscopy*. SPIE BiOS, San Francisco, California, USA.

He, F., Liang, N.-N., Mu, L., Pan, Q.-H., Wang, J., Reeves, M. J., & Duan, C.-Q. (2012). Anthocyanins and their variation in red wines I. Monomeric anthocyanins and their color expression. *Molecules, 17*(2), 1571-1601. https://doi.org/10.3390/molecules17021571

Heymann, H., & Noble, A. C. (1987). Descriptive analysis of commercial Cabernet-Sauvignon wines from California. *American Journal of Enology and Viticulture, 38*(1), 41-44. https://www.ajevonline.org/content/ajev/38/1/41.full.pdf

Hopfer, H., Buffon, P. A., Ebeler, S. E., & Heymann, H. (2013). The combined effects of storage temperature and packaging on the sensory, chemical, and physical properties of a Cabernet-Sauvignon wine. *Journal of Agricultural and Food Chemistry, 61*(13), 3320-3334. https://doi.org/10.1021/jf3051736

Hopfer, H., Nelson, J., Ebeler, S. E., & Heymann, H. (2015). Correlating wine quality indicators to chemical and sensory measurements. *Molecules, 20*(5), 8453-8483. https://doi.org/10.3390/molecules20058453

Iland, P., Bruer, N., Edwards, G., Weeks, S., & Wilkes, E. (2004). Chemical Analysis of Grapes and Wine: Techniques and Concepts. Patrick Iland Wine Promotions Pty Ltd.

Janik, L. J., Cozzolino, D., Dambergs, R., Cynkar, W., & Gishen, M. (2007). The prediction of total anthocyanin concentration in red-grape homogenates using visible-near-infrared spectroscopy and artificial neural networks. *Analytica Chimica Acta, 594*(1), 107-118. https://doi.org/10.1016/j.aca.2007.05.019

Karoui, R., & Blecker, C. (2011). Fluorescence spectroscopy measurement for quality assessment of food systems—A review. *Food and Bioprocess Technology, 4*(3), 364-386. https://doi.org/10.1007/s11947-010-0370-0

Kuhn, N., Guan, L., Dai, Z. W., Wu, B.-H., Lauvergeat, V., Gomès, E., Li, S.-H., Godoy, F., Arce-Johnson, P., & Delrot, S. (2013). Berry ripening: recently heard through the grapevine. *Journal of Experimental Botany, 65*(16), 4543-4559. https://doi.org/10.1093/jxb/ert395

Kustos, M., Gambetta, J. M., Jeffery, D. W., Heymann, H., Goodman, S., & Bastian, S. E. P. (2020). A matter of place: Sensory and chemical characterisation of fine Australian Chardonnay and Shiraz wines of provenance. *Food Research International, 130*, 108903. https://doi.org/10.1016/j.foodres.2019.108903

Li, S., Bindon, K., Bastian, S. E. P., Jiranek, V., & Wilkinson, K. L. (2017). Use of winemaking supplements to modify the composition and sensory properties of Shiraz wine. *Journal of Agricultural and Food Chemistry, 65*(7), 1353-1364. https://doi.org/10.1021/acs.jafc.6b04505

Mandrile, L., Zeppa, G., Giovannozzi, A. M., & Rossi, A. M. (2016). Controlling protected designation of origin of wine by Raman spectroscopy. *Food Chemistry, 211*, 260-267. https://doi.org/10.1016/j.foodchem.2016.05.011

Mezei, L. V., Johnson, T. E., Goodman, S., Collins, C., & Bastian, S. E. P. (2021). Meeting the demands of climate change: Australian consumer acceptance and sensory profiling of red wines produced from non-traditional red grape varieties. *OENO One, 55*(2), 29-46. https://doi.org/10.20870/oeno-one.2021.55.2.4571

Murphy, K., Stedmon, C., Graeber, D., & Bro, R. (2013). Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Analytical Methods, 5*(23), 6557-6566. https://doi.org/10.1039/C3AY41160E

Nguyen, A. N. H., Johnson, T. E., Jeffery, D. W., Capone, D. L., Danner, L., & Bastian, S. E. P. (2020). Sensory and chemical drivers of wine consumers' preference for a new Shiraz wine product containing *Ganoderma lucidum* extract as a novel ingredient. *Foods, 9*(2). https://doi.org/10.3390/foods9020224

Niimi, J., Tomic, O., Næs, T., Jeffery, D. W., Bastian, S. E. P., & Boss, P. K. (2018). Application of sequential and orthogonalised-partial least squares (SO-PLS) regression to predict sensory properties of Cabernet-Sauvignon wines from grape chemical composition. *Food Chemistry, 256*, 195-202. https://doi.org/10.1016/j.foodchem.2018.02.120

Paissoni, M. A., Waffo-Teguo, P., Ma, W., Jourdes, M., Rolle, L., & Teissedre, P. L. (2018). Chemical and sensorial investigation of in-mouth sensory properties of grape anthocyanins. *Scientific Reports, 8*(1), 17098. https://doi.org/10.1038/s41598-018-35355-x

Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses, 27*(8), 747-755. https://doi.org/10.1093/chemse/27.8.747

Pawlus, A. D., Waffo-Téguo, P., Shaver, J., & Mérillon, J.-M. (2012). Stilbenoid chemistry from wine and the genus *Vitis*, a review. *OENO One, 46*(2), 57-111. https://doi.org/10.20870/oeno-one.2012.46.2.1512

Picard, M., de Revel, G., & Marchand, S. (2017). First identification of three *p*-menthane lactones and their potential precursor, menthofuran, in red wines. *Food Chemistry, 217*, 294-302. https://doi.org/10.1016/j.foodchem.2016.08.070

Picard, M., Lytra, G., Tempere, S., Barbe, J.-C., de Revel, G., & Marchand, S. (2016a). Identification of piperitone as an aroma compound contributing to the positive mint nuances perceived in aged red Bordeaux wines. *Journal of Agricultural and Food Chemistry, 64*(2), 451-460. https://doi.org/10.1021/acs.jafc.5b04869

Picard, M., Tempere, S., de Revel, G., & Marchand, S. (2016b). Piperitone profiling in fine red Bordeaux wines: Geographical influences in the Bordeaux region and enantiomeric distribution. J*ournal of Agricultural and Food Chemistry, 64*(40), 7576-7584. https://doi.org.10.1021/acs.jafc.6b02835

Picard, M., Tempere, S., de Revel, G., & Marchand, S. (2015). A sensory study of the ageing bouquet of red Bordeaux wines: A three-step approach for exploring a complex olfactory concept. *Food Quality and Preference, 42*, 110-122. https://doi.org/10.1016/j.foodqual.2015.01.014

Ranaweera, R. K. R., Capone, D. L., Bastian, S. E. P., Cozzolino, D., & Jeffery, D. W. (2021a). A review of wine authentication using spectroscopic approaches in combination with chemometrics. *Molecules, 26*(14). https://doi.org/10.3390/molecules26144334

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021b). Authentication of the geographical origin of Australian Cabernet-Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. *Food Chemistry, 335*, 127592. https://doi.org/10.1016/j.foodchem.2020.127592

Ranaweera, R. K. R., Gilmore, A. M., Capone, D. L., Bastian, S. E. P., & Jeffery, D. W. (2021c). Spectrofluorometric analysis combined with machine learning for geographical and varietal authentication, and prediction of phenolic compound concentrations in red wine. *Food Chemistry 361*, 130149. https://doi.org/10.1016/j.foodchem.2021.130149

Riovanto, R., Cynkar, W. U., Berzaghi, P., & Cozzolino, D. (2011). Discrimination between Shiraz wines from different Australian regions: the role of spectroscopy and chemometrics. J*ournal of Agricultural and Food Chemistry, 59*(18), 10356-10360. https://doi.org/10.1021/jf202578f

Ristic, R., Bindon, K., Francis, L. I., Herderich, M. J., & Iland, P. G. (2010). Flavonoids and C_{13}-norisoprenoids in *Vitis vinifera* L. cv. Shiraz: relationships between grape and wine composition, wine colour and wine sensory properties. *Australian Journal of Grape and Wine Research, 16*(3), 369-388. https://doi.org/10.1111/j.1755-0238.2010.00099.x

Robinson, A. L., Adams, D. O., Boss, P. K., Heymann, H., Solomon, P. S., & Trengove, R. D. (2011). The relationship between sensory attributes and wine composition for Australian Cabernet-Sauvignon wines. *Australian Journal of Grape and Wine Research, 17*(3), 327-340. https://doi.org/10.1111/j.1755-0238.2011.00155.x

Sádecká, J., & Jakubíková, M. (2020). Varietal classification of white wines by fluorescence spectroscopy. *Journal of Food Science and Technology, 57*(7), 2545-2553. https://doi.org/10.1007/s13197-020-04291-y

Schueuermann, C., Silcock, P., & Bremer, P. (2018). Front-face fluorescence spectroscopy in combination with parallel factor analysis for profiling of clonal and vineyard site differences in commercially produced Pinot Noir grape juices and wines. *Journal of Food Composition and Analysis, 66*, 30-38. https://doi.org/10.1016/j.jfca.2017.11.005

Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., Danner, L., & Jeffery, D. W. (2019). Using content analysis to characterise the sensory typicity and quality judgements of Australian Cabernet-Sauvignon wines. *Foods, 8*(12), 691. https://doi.org/10.3390/foods8120691

Souza Gonzaga, L., Capone, D. L., Bastian, S. E. P., Danner, L., & Jeffery, D. W. (2020). Sensory typicity of regional Australian Cabernet-Sauvignon wines according to expert evaluations and descriptive analysis. *Food Research International, 138*, Part A, 109760. https://doi.org/10.1016/j.foodres.2020.109760

Suciu, R.-C., Zarbo, L., Guyon, F., & Magdas, D. A. (2019). Application of fluorescence spectroscopy using classical right angle technique in white wines classification. *Scientific Reports, 9*(1), 18250. https://doi.org/10.1038/s41598-019-54697-8

Tan, J., Li, R., Jiang, Z.-T., Zhang, Y., Hou, Y.-M., Wang, Y.-R., Wu, X., & Gong, L. (2016). Geographical classification of Chinese Cabernet-Sauvignon wines by data fusion of ultraviolet–visible and synchronous fluorescence spectroscopies: the combined use of multiple wavelength differences. *Australian Journal of Grape and Wine Research, 22*(3), 358-365. https://doi.org/10.1111/ajgw.12214

Vidal, S., Francis, L., Williams, P., Kwiatkowski, M., Gawel, R., Cheynier, V., & Waters, E. (2004). The mouth-feel properties of polysaccharides and anthocyanins in a wine like medium. *Food Chemistry, 85*(4), 519-525. https://doi.org/10.1016/S0308-8146(03)00084-0

Waterhouse, A. L., Sacks, G. L., & Jeffery, D. W. (2016a). *Flavan-3-ols and Condensed Tannin. In Understanding Wine Chemistry* (pp. 117-126). Wiley. https://doi.org/10.1002/9781118730720.ch14

Waterhouse, A. L., Sacks, G. L., & Jeffery, D. W. (2016b). *Introduction to Phenolics*. In Understanding Wine Chemistry (pp. 99-104). Wiley. https://doi.org/10.1002/9781118730720.ch11

SUPPLEMENTARY DATA

Souza Gonzaga, L., Bastian, S., Capone, D., Ranaweera, R., & Jeffery, D. (2021). Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *OENO One, 55*(4)
https://doi.org/10.20870/oeno-one.2021.55.4.4805

TABLE S1. Wine sample code with winery, blend components, oak contact information, and basic analytical measurements. Values for free and total SO_2, pH and TA are the means of duplicate measurements, and those for ethanol are from triplicate measurements.

Sample code	Winery	Composition of blend	Type of oak contact	pH	TA (gL^{-1})	Free SO_2 (mgL^{-1})	Total SO_2 (mgL^{-1})	Ethanol % v/v
CS1	A	100 % Cabernet-Sauvignon	Initial barrel maturation[1]	3.61	7.0	22.0	42.0	14.7
CS2	A	79 % Cabernet-Sauvignon, 21 % Petit Verdot	Initial barrel maturation	3.55	7.2	30.4	52.4	15.2
CS3	A	86 % Cabernet-Sauvignon, 14 % Petit Verdot	Initial barrel maturation	3.58	7.2	26.0	48.0	14.8
CS4	A	100 % Cabernet-Sauvignon	Initial barrel maturation	3.54	7.4	33.6	70.8	15.3
CS5	B	100 % Cabernet-Sauvignon	No contact	3.50	7.0	0.4	0.4	15.1
CS6	B	100 % Cabernet-Sauvignon	Initial barrel maturation	3.48	7.5	1.2	8.0	15.3
CS7	C	100 % Cabernet-Sauvignon	Initial barrel maturation	3.87	6.5	4.4	9.2	14.7
CS8	C	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.40	7.0	13.6	20.8	15.2
CS9	D	100 % Cabernet-Sauvignon	Initial barrel maturation	3.67	5.9	13.2	21.2	14.7
CS10	D	100 % Cabernet-Sauvignon	Initial barrel maturation	3.55	5.9	20.0	33.2	13.7
CS11	D	100 % Cabernet-Sauvignon	Initial barrel maturation	3.63	5.9	21.6	26.4	14.9
CS12	E	100 % Cabernet-Sauvignon	Initial barrel maturation	3.73	5.6	29.2	46.8	14.5
CS13	E	100 % Cabernet-Sauvignon	Initial barrel maturation	3.82	5.6	23.6	37.2	14.3
CS14	F	100 % Cabernet-Sauvignon	Initial barrel maturation	3.46	6.1	20.0	39.2	12.9
CS15	G	100 % Cabernet-Sauvignon	Initial barrel maturation	3.55	6.1	29.2	46.4	14.6
CS16	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.73	7.0	20.0	39.6	14.4
CS17	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.53	6.8	16.8	26.8	13.7
CS18	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.60	7.2	23.6	39.6	14.3
CS19	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.58	6.9	23.2	45.2	13.7
CS20	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.59	6.7	22.0	40.8	13.8
CS21	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.63	6.5	14.8	26.8	15.1
CS22	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.74	6.4	18.8	26.8	14.1
CS23	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.55	6.6	18.0	31.6	15.1
CS24	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.58	6.5	23.2	37.2	14.2
CS25	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.50	6.9	16.0	25.6	14.1
CS26	H	100 % Cabernet-Sauvignon	Oak chips during fermentation	3.67	7.44	20.0	34.0	13.8

[1] Initial barrel maturation was ≤ 5 months

SUPPLEMENTARY DATA

Souza Gonzaga, L., Bastian, S., Capone, D., Ranaweera, R., & Jeffery, D. (2021).
Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *OENO One, 55*(4)
https://doi.org/10.20870/oeno-one.2021.55.4.4805

TABLE S2. Aroma/flavour, taste, and mouthfeel attributes used for the RATA evaluation with respective description available to panellists during tasting.

Attribute	Description
Aroma/flavour	
Dark fruits	blackcurrant, blackberry, mulberry, plum, blueberry
Red fruits	strawberry, raspberry
Confectionery	lollies
Jammy	preserved or cooked fruit
Dried fruits	raisin, prune, sultana
Spices	clove, cinnamon, nutmeg, mixed spice, mulled wine, liquorice, anise
Savoury	meaty, soy sauce, black olives, salami
Green	green pepper, capsicum, asparagus, green beans
Tobacco	smoky, cigar
Earthy	mushroom, truffle, forest floor
Minty	spearmint, mint, peppermint
Eucalyptus	menthol, VapoRub ointment
Violets	blue flowers
Oaky	woody, pencil shaving, toasty, cedar, coconut
Cooked Vegetables	pickles, cabbage, canned beans or mixed veggies
Grassy	stalky, leafy, cut grass, herbaceous, tomato leaf
Floral	roses, perfume, musk
Vanilla/Chocolate	vanilla and chocolate
Dried herbs	oregano
Barnyard	band-aid, medicinal, horse-like, stables
Cherry cola	sarsaparilla
Brine	oyster shell, sea spray, sardine oil
Yeasty	bread, doughy
Taste	
Bitterness	bitter taste
Acidity	tart, sour taste
Sweetness	sweet taste
Mouthfeel	
Body	weight of the wine when you swirl it in your mouth, thickness, mouth feeling
Astringency	drying, roughing and puckering sensation
Alcohol	warmth or burning sensation
Length of flavour	how persistent was the wine flavour in the mouth after expectoration

SUPPLEMENTARY DATA

Souza Gonzaga, L., Bastian, S., Capone, D., Ranaweera, R., & Jeffery, D. (2021). Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *OENO One, 55*(4)
https://doi.org/10.20870/oeno-one.2021.55.4.4805

TABLE S3. Mean values of each sensory attribute along with p-values according to ANOVA (significant differences with α = 0.1 are shown in bold) for Cabernet-Sauvignon wines from Coonawarra (n = 26; refer to Table S1 for wine codes).

Attributes	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	CS9	CS10	CS11	CS12	CS13	CS14	CS15	CS16	CS17	CS18	CS19	CS20	CS21	CS22	CS23	CS24	CS25	CS26	p-value
A-Barnyard	**0.83**	**0.48**	**0.87**	**1.10**	**1.07**	**0.67**	**0.82**	**0.68**	**1.42**	**1.07**	**0.70**	**1.03**	**0.85**	**0.73**	**0.75**	**0.73**	**0.65**	**0.77**	**1.40**	**1.23**	**0.92**	**0.70**	**0.43**	**0.80**	**0.95**	**0.78**	**0.004**
A-Brine	0.67	0.55	0.43	0.38	0.37	0.52	0.43	0.48	0.45	0.68	0.77	0.58	0.43	0.47	0.77	0.48	0.28	0.48	0.85	0.80	0.55	0.52	0.48	0.52	0.53	0.18	0.233
A-Cherry cola	0.90	1.00	0.85	0.82	1.53	1.62	1.12	1.37	0.70	1.07	1.18	1.02	1.03	1.17	0.90	1.07	1.08	1.08	0.93	0.98	1.10	1.07	1.18	1.13	1.17	1.20	0.440
A-Confectionery	1.22	1.50	1.38	1.23	1.85	1.90	1.48	1.67	1.22	1.25	1.47	1.57	1.28	1.33	1.58	1.55	1.22	1.15	1.70	1.53	1.15	1.43	1.57	1.92	1.15	1.75	0.268
A-Cooked Vegetables	**0.93**	**0.57**	**0.93**	**0.67**	**0.68**	**0.47**	**0.57**	**0.58**	**1.12**	**0.80**	**0.92**	**0.82**	**0.88**	**0.55**	**1.02**	**0.73**	**0.60**	**0.67**	**0.60**	**0.53**	**1.22**	**1.15**	**0.78**	**1.17**	**1.07**	**0.77**	**0.021**
A-Dark fruits	**3.63**	**3.32**	**3.75**	**3.43**	**3.05**	**3.55**	**3.07**	**3.95**	**3.02**	**3.17**	**3.65**	**3.52**	**3.32**	**3.35**	**3.37**	**3.43**	**3.25**	**3.17**	**2.83**	**2.88**	**3.38**	**2.90**	**3.47**	**2.97**	**3.17**	**3.00**	**0.063**
A-Dried fruits	1.92	2.00	2.30	1.98	2.25	2.10	1.68	2.30	1.57	2.08	2.15	2.05	1.75	2.30	2.02	1.73	1.90	2.15	1.77	1.72	2.17	1.75	2.28	1.92	2.02	1.88	0.692
A-Dried herbs	1.05	0.87	1.02	0.95	0.82	0.58	1.15	0.90	0.82	0.62	0.92	1.03	0.73	0.73	0.72	1.02	0.75	0.78	1.22	1.00	0.77	1.15	0.90	0.68	1.10	0.75	0.453
A-Earthy	**2.10**	**1.33**	**1.93**	**1.47**	**1.20**	**1.00**	**1.12**	**1.67**	**1.20**	**1.22**	**1.58**	**1.73**	**1.10**	**1.08**	**1.97**	**1.27**	**1.22**	**1.38**	**1.57**	**1.27**	**1.35**	**1.35**	**1.37**	**1.45**	**1.75**	**1.25**	**0.004**
A-Eucalyptus	1.07	1.07	1.17	1.07	1.32	1.13	1.58	1.02	0.88	0.78	0.78	0.85	0.82	0.97	1.00	0.78	0.68	0.77	0.97	1.12	0.67	0.75	0.90	0.70	0.92	0.85	0.115
A-Floral	1.48	1.27	1.97	1.40	1.77	2.07	1.65	1.40	1.38	1.53	1.30	1.42	1.28	1.02	1.37	1.28	1.38	1.37	1.30	1.62	1.30	1.17	1.65	1.42	1.62	1.52	0.301
A-Grassy	1.38	1.10	1.38	1.18	1.03	0.98	1.05	1.53	1.23	1.00	1.75	1.13	1.22	1.02	1.17	1.02	0.97	1.20	1.75	1.08	1.40	1.35	1.23	1.35	1.40	1.05	0.198
A-Green	1.08	1.02	1.27	1.27	1.28	0.88	1.40	1.28	0.83	1.10	1.10	1.07	0.82	1.35	1.10	1.27	1.03	1.17	1.17	1.25	1.40	1.48	1.18	1.33	1.15	1.20	0.754
A-Jammy	2.47	1.82	1.80	1.97	2.23	2.37	1.90	2.52	1.70	2.05	2.10	2.18	1.95	1.85	2.40	2.07	1.73	2.08	1.40	1.73	2.18	2.02	2.37	2.12	2.07	1.87	0.140
A-Minty	**0.87**	**1.23**	**1.52**	**1.28**	**1.08**	**1.35**	**1.18**	**1.13**	**0.80**	**0.92**	**1.28**	**0.88**	**0.75**	**1.20**	**0.97**	**0.92**	**0.87**	**0.68**	**1.17**	**1.03**	**0.95**	**0.87**	**0.90**	**1.03**	**1.47**	**1.62**	**0.010**
A-Oaky	**1.88**	**1.83**	**1.93**	**1.88**	**1.72**	**1.60**	**1.38**	**2.25**	**1.70**	**1.43**	**1.92**	**2.20**	**1.60**	**1.90**	**1.57**	**1.80**	**1.48**	**1.75**	**1.65**	**1.25**	**1.62**	**1.40**	**1.65**	**1.53**	**1.98**	**1.30**	**0.069**
A-Red fruits	2.88	2.13	2.87	2.35	2.47	2.52	2.20	2.83	2.43	2.45	2.60	2.82	2.37	2.82	2.45	2.08	2.70	2.57	2.02	2.50	1.90	2.17	2.82	2.47	2.28	2.37	0.104
A-Savoury	**1.72**	**1.38**	**1.50**	**1.25**	**1.10**	**0.60**	**0.85**	**1.30**	**1.23**	**1.50**	**1.30**	**1.30**	**0.97**	**0.97**	**1.35**	**1.20**	**0.85**	**0.93**	**1.30**	**0.82**	**1.43**	**1.25**	**1.30**	**1.32**	**1.22**	**0.87**	**0.013**
A-Spices	2.22	2.37	2.23	2.03	2.12	1.83	2.05	1.95	1.33	1.68	2.12	2.32	1.92	2.18	1.97	1.50	1.58	2.00	1.75	1.82	1.95	1.85	1.93	1.93	1.83	1.87	0.237
A-Tobacco	**1.18**	**0.93**	**1.20**	**0.78**	**0.78**	**0.53**	**0.83**	**1.20**	**0.75**	**0.67**	**1.28**	**1.07**	**0.77**	**0.88**	**0.72**	**0.65**	**0.70**	**1.05**	**1.13**	**0.92**	**1.03**	**1.10**	**0.93**	**0.90**	**1.17**	**0.48**	**0.024**
A-Vanilla/Chocolate	0.87	1.18	1.02	0.92	1.27	1.07	0.97	1.05	0.65	0.70	0.77	0.98	0.62	0.78	0.85	0.78	1.03	0.92	0.65	0.72	0.83	0.68	1.22	0.68	0.80	0.72	0.347
A-Violets	0.62	0.73	1.10	1.00	0.88	1.28	1.20	0.73	0.83	0.85	0.87	0.88	0.80	0.92	0.73	1.08	1.22	0.90	0.90	0.92	0.67	0.95	0.88	0.87	1.17	1.08	0.581
A-Yeasty	0.57	0.60	0.55	0.43	0.58	0.67	0.35	0.75	0.35	0.42	0.63	0.75	0.47	0.37	0.88	0.58	0.62	0.45	0.57	0.28	0.57	0.48	0.63	0.57	0.48	0.45	0.387
F-Barnyard	**0.57**	**0.58**	**0.82**	**0.62**	**0.92**	**0.52**	**0.57**	**0.57**	**0.95**	**0.65**	**0.50**	**0.67**	**0.63**	**0.88**	**0.68**	**0.48**	**0.47**	**0.65**	**1.60**	**0.95**	**0.65**	**0.85**	**0.37**	**0.62**	**0.87**	**0.48**	**< 0.0001**
F-Brine	0.60	0.45	0.48	0.65	0.65	0.53	0.45	0.88	0.58	0.38	0.48	0.57	0.37	0.43	0.95	0.40	0.53	0.47	0.82	0.68	0.40	0.48	0.43	0.62	0.57	0.38	0.151
F-Cherry cola	**0.87**	**0.85**	**0.88**	**1.03**	**1.35**	**1.65**	**1.32**	**1.17**	**0.73**	**0.85**	**1.12**	**0.65**	**0.77**	**1.10**	**0.80**	**0.82**	**1.30**	**0.98**	**0.77**	**0.93**	**1.10**	**1.45**	**1.10**	**1.25**	**1.25**	**1.18**	**0.007**
F-Confectionery	1.33	1.30	1.27	1.20	1.60	2.08	1.28	1.43	0.97	1.23	1.48	1.22	1.37	1.13	1.22	1.25	1.65	1.37	1.02	1.27	1.05	1.43	1.48	1.57	1.27	1.27	0.212
F-Cooked Vegetables	**0.77**	**0.40**	**0.95**	**0.60**	**0.47**	**0.53**	**0.60**	**0.57**	**0.88**	**0.45**	**0.92**	**0.55**	**0.62**	**0.55**	**0.83**	**0.75**	**0.65**	**0.90**	**0.68**	**0.58**	**0.95**	**1.07**	**0.50**	**1.15**	**0.93**	**0.55**	**0.026**
F-Dark fruits	**3.78**	**3.53**	**3.47**	**3.70**	**3.52**	**3.92**	**3.50**	**3.97**	**3.38**	**3.22**	**4.02**	**3.63**	**3.27**	**3.18**	**3.47**	**3.75**	**3.33**	**3.45**	**2.88**	**3.13**	**3.85**	**3.25**	**3.60**	**3.25**	**3.65**	**2.93**	**0.008**
F-Dried fruits	1.90	1.80	1.72	1.92	2.12	1.82	1.78	2.17	1.73	2.10	1.90	2.23	1.97	1.77	2.00	1.85	1.60	1.93	1.52	1.75	2.13	2.02	2.23	1.93	2.33	2.12	0.720
F-Dried herbs	0.90	0.93	0.83	1.02	0.77	0.97	0.87	0.95	0.98	0.90	0.90	0.95	0.88	0.67	1.00	1.10	0.85	0.55	0.98	1.05	0.75	0.98	1.02	0.82	1.10	0.82	0.976
F-Earthy	**1.80**	**1.63**	**1.62**	**1.53**	**1.62**	**1.20**	**0.97**	**1.75**	**1.13**	**0.98**	**1.77**	**1.80**	**1.37**	**1.38**	**1.90**	**0.90**	**0.92**	**1.45**	**1.55**	**1.43**	**1.43**	**1.25**	**1.50**	**1.38**	**1.72**	**1.32**	**0.001**
F-Eucalyptus	0.90	1.12	1.13	0.87	1.13	1.03	0.95	0.93	0.60	0.67	0.77	0.67	0.67	0.90	0.63	0.68	0.52	0.80	0.90	0.80	0.68	0.58	0.88	0.73	0.88	0.80	0.332
F-Floral	1.22	0.98	1.20	0.93	1.45	1.53	1.28	1.22	0.97	1.40	0.88	1.13	1.10	1.00	1.07	1.15	1.25	1.15	0.92	0.88	0.95	1.12	1.58	1.03	1.07	1.30	0.431
F-Grassy	1.30	1.12	1.20	1.43	1.05	0.98	1.07	1.32	0.97	1.07	1.18	1.23	0.82	0.88	1.33	1.05	0.77	1.10	1.12	1.28	1.20	0.95	1.28	1.17	1.45	0.88	0.496
F-Green	1.15	1.12	1.08	1.28	1.33	0.82	0.87	1.03	1.15	1.00	1.15	1.00	1.17	1.02	1.23	1.27	1.12	1.13	1.27	1.32	1.20	1.32	1.20	1.37	1.33	1.18	0.938
F-Jammy	2.33	2.00	1.82	1.90	2.13	2.13	2.15	2.22	1.48	1.75	2.20	1.88	1.53	1.82	2.15	1.82	1.78	1.77	1.88	1.77	2.45	2.13	2.20	2.30	1.78	1.72	0.161

SUPPLEMENTARY DATA

Souza Gonzaga, L., Bastian, S., Capone, D., Ranaweera, R., & Jeffery, D. (2021). Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *OENO One, 55*(4)
https://doi.org/10.20870/oeno-one.2021.55.4.4805

TABLE S3. CONTD.

Attributes	CS1	CS2	CS3	CS4	CS5	CS6	CS7	CS8	CS9	CS10	CS11	CS12	CS13	CS14	CS15	CS16	CS17	CS18	CS19	CS20	CS21	CS22	CS23	CS24	CS25	CS26	p-value
F-Minty	**0.78**	**0.97**	**1.32**	**1.28**	**1.22**	**1.10**	**1.27**	**1.12**	**0.80**	**0.75**	**1.02**	**0.67**	**0.68**	**1.05**	**0.82**	**0.83**	**0.85**	**0.68**	**0.85**	**0.68**	**0.97**	**0.85**	**0.85**	**0.85**	**1.10**	**1.10**	**0.063**
F-Oaky	2.15	2.45	2.22	2.20	2.07	2.10	1.50	2.35	1.97	1.88	2.33	2.13	1.92	1.70	2.18	1.80	1.45	2.13	1.87	1.60	1.95	1.83	2.10	1.90	2.13	1.95	0.127
F-Red fruits	2.70	2.40	2.82	2.53	2.52	2.43	2.67	2.73	2.20	2.45	2.53	2.30	2.27	2.73	2.62	2.13	3.22	2.50	1.93	2.50	2.38	2.28	2.77	2.57	2.67	2.25	0.165
F-Savoury	1.52	1.38	1.47	1.63	1.08	1.02	1.12	1.45	1.42	1.40	1.38	1.25	1.05	1.08	1.52	1.08	0.77	1.43	1.30	1.32	1.40	1.32	1.57	1.55	1.63	0.93	0.141
F-Spices	2.40	2.25	2.30	2.35	2.30	2.12	2.42	2.45	1.87	1.85	2.42	2.43	1.55	1.98	1.90	2.10	1.80	2.13	2.20	1.90	2.48	2.00	2.27	2.10	1.73	2.10	0.105
F-Tobacco	**1.13**	**0.92**	**1.05**	**1.20**	**0.97**	**0.72**	**0.72**	**1.13**	**0.60**	**0.62**	**1.12**	**0.88**	**0.72**	**0.70**	**0.75**	**0.73**	**0.50**	**0.97**	**1.05**	**0.60**	**1.08**	**0.70**	**0.90**	**1.05**	**1.15**	**0.82**	**0.099**
F-Vanilla/Chocolate	**0.73**	**0.78**	**0.83**	**0.68**	**0.70**	**0.90**	**0.98**	**1.18**	**0.88**	**1.00**	**0.73**	**0.60**	**0.80**	**0.43**	**0.50**	**0.73**	**1.33**	**0.63**	**0.40**	**0.57**	**0.80**	**0.82**	**0.98**	**0.47**	**0.75**	**0.68**	**0.010**
F-Violets	0.55	0.82	0.78	0.65	1.02	1.03	1.10	0.77	0.53	0.75	0.63	0.62	0.68	0.75	0.60	0.60	1.07	0.63	0.73	0.93	0.62	0.70	0.93	0.85	0.85	1.03	0.248
F-Yeasty	0.42	0.37	0.43	0.27	0.48	0.30	0.65	0.50	0.60	0.45	0.47	0.48	0.20	0.58	0.50	0.47	0.32	0.57	1.00	0.75	0.38	0.48	0.42	0.45	0.47	0.37	0.103
T-Acidity	**3.97**	**3.48**	**3.80**	**3.52**	**3.75**	**3.62**	**3.15**	**4.17**	**3.20**	**3.38**	**3.73**	**4.18**	**3.33**	**3.58**	**3.75**	**3.45**	**3.20**	**3.57**	**3.75**	**3.73**	**3.17**	**3.00**	**3.83**	**3.23**	**3.68**	**3.62**	**< 0.0001**
T-Bitterness	**2.80**	**2.83**	**2.73**	**2.80**	**2.85**	**2.78**	**2.02**	**2.72**	**2.68**	**2.55**	**2.77**	**3.07**	**2.63**	**2.87**	**2.90**	**2.58**	**2.08**	**2.68**	**3.15**	**2.55**	**2.52**	**2.55**	**2.77**	**2.25**	**2.87**	**2.60**	**0.029**
T-Sweetness	**2.23**	**2.15**	**2.22**	**2.13**	**2.48**	**2.43**	**2.75**	**2.15**	**2.23**	**2.35**	**2.63**	**1.82**	**2.07**	**2.07**	**2.25**	**2.22**	**2.70**	**2.18**	**1.77**	**1.95**	**2.65**	**2.52**	**2.30**	**2.57**	**2.18**	**2.03**	**0.004**
MF-Alcohol	3.87	3.78	3.88	3.92	3.93	4.17	3.58	3.80	4.00	3.75	3.90	3.70	3.75	3.77	3.80	3.92	3.67	3.62	3.50	3.67	3.85	3.78	3.75	3.53	3.62	3.60	0.447
MF-Astringency	**3.85**	**4.10**	**4.48**	**4.05**	**4.15**	**4.17**	**3.37**	**4.17**	**4.20**	**3.67**	**4.10**	**4.93**	**4.32**	**3.95**	**4.55**	**4.05**	**3.60**	**4.02**	**4.82**	**4.05**	**3.92**	**3.68**	**4.23**	**3.75**	**4.10**	**4.65**	**< 0.0001**
MF-Body	3.65	3.77	4.02	3.95	4.17	3.97	3.82	3.73	3.77	3.58	3.80	3.83	3.62	3.72	3.80	3.97	3.57	3.70	3.75	3.57	3.92	3.70	3.92	3.55	3.68	3.73	0.272
Length of flavour	4.35	4.45	4.55	4.40	4.50	4.62	4.27	4.35	4.27	4.08	4.47	4.30	4.27	4.12	4.32	4.28	4.07	4.05	4.33	4.12	4.42	4.33	4.47	4.03	4.38	4.27	0.115

SUPPLEMENTARY DATA

Souza Gonzaga, L., Bastian, S., Capone, D., Ranaweera, R., & Jeffery, D. (2021). Modelling Cabernet-Sauvignon wine sensory traits from spectrofluorometric data. *OENO One, 55*(4)
https://doi.org/10.20870/oeno-one.2021.55.4.4805

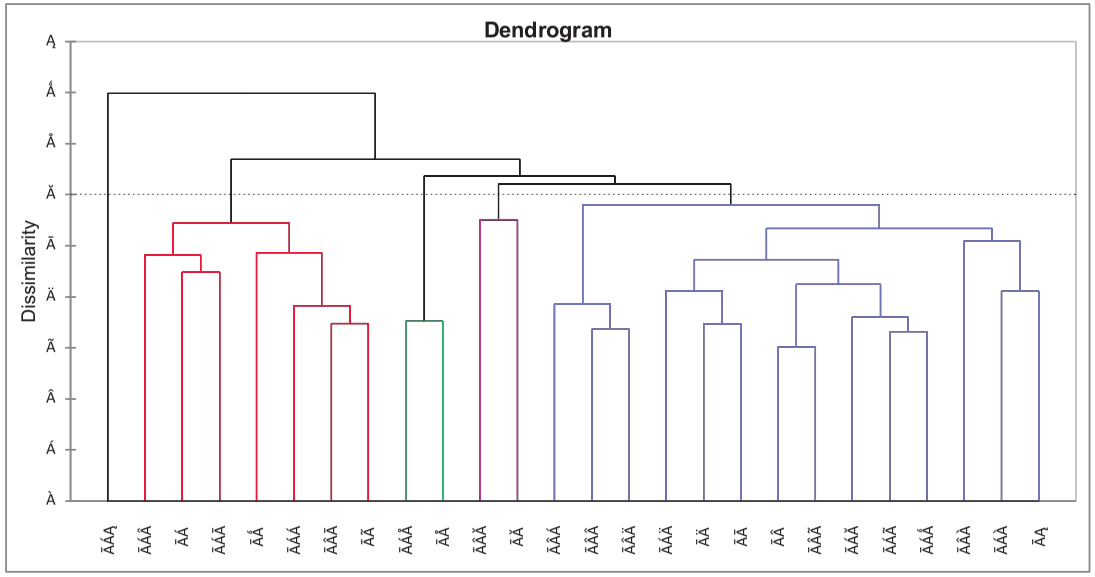


FIGURE S1. Agglomerative hierarchical clustering dendrogram plot of the five clusters segregating the Coonawarra Cabernet-Sauvignon wines (n = 26, refer to Table S1 for wine codes) evaluated with rate-all-that-apply (significantly different attributes as presented in Table S3, α = 0.1). The dotted line shows the truncation according to the dissimilarity on the y-axis.

Appendix B

Conference Proceeding

Application of fluorescence spectroscopy with multivariate analysis for authentication of shiraz wines from different regions

Ranaweera K.R. Ranaweera, Adam M. Gilmore, Dimitra L. Capone, Susan E.P. Bastian, David W. Jeffery (2020). Authenticating the geographical origin of wine using fluorescence spectroscopy and machine learning, Paper presented at the XIIIth International Terroir Congress, Adelaide, Australia.

XIIIth International Terroir Congress
17–18 November 2020
Virtual Congress | Adelaide, Australia

APPLICATION OF FLUORESCENCE SPECTROSCOPY WITH MULTIVARIATE ANALYSIS FOR AUTHENTICATION OF SHIRAZ WINES FROM DIFFERENT REGIONS

R.K.R. Ranaweera[a], A. M. Gilmore[b], D.L. Capone[a, c], S.E.P. Bastian[a,c], D.W Jeffery[a, c]*

[a]Department of Wine and Food Science, The University of Adelaide, South Australia, Australia
[b]HORIBA Instruments Inc., 20 Knightsbridge Rd., Piscataway, NJ 08854, United States
[c]Australian Research Council Training Centre for Innovative Wine Production, The University of Adelaide

*Corresponding author: david.jeffery@adelaide.edu.au

Abstract

Aim: To investigate the possibility of utilising simultaneous measurements of absorbance-transmittance and fluorescence excitation-emission matrix (A-TEEM) combined with chemometrics, as a robust method that gives rapid results for classification of wines from different regions of South Australia according to their Geographical Indication (GI), and to gain insight into the effect of terroir on inter regional variation.

Methods and Results: Additionally, to obtaining various colour parameters, the A-TEEM technique enables the "fingerprint" of wine samples to be attained in response to the presence of fluorophoric compounds. This is accomplished by recording a three-dimensional excitation-emission matrix (EEM) over multiple excitation and emission wavelengths, which can then be analysed using multivariate statistical modelling to classify wines. Shiraz wine samples (n = 134) from six different GIs of South Australia (Barossa Valley, Clare Valley, Eden Valley, Langhorne Creek, McLaren Vale, and Riverland) were analysed and absorbance spectra, hue, intensity, CIE L*a*b, CIE 1931, and EEMs were recorded for each sample. EEM data were evaluated according to the cross-validation model built with extreme gradient boost discriminant analysis (XGBDA) using score probability to assess the accuracy of classification according to the region of origin. Preliminary results have shown a high prediction ability and the data extracted from A-TEEM could be used to investigate phenolics as potential chemical markers that may provide effective regional discrimination.

Conclusions: The molecular fingerprinting capability and sensitivity of EEM in conjunction with multivariate statistical analysis of the fluorescence data using the XGBDA algorithm provided sufficient chemical/spectral information to facilitate accurate classification of Shiraz wines according to the region of origin. A-TEEM coupled with XGBDA modelling appears to be a promising tool for wine authentication according to its geographical origin.

Significance and Impact of the Study: Having tangible evidence that Australian fine wines may be discriminated on the basis of geographical origin, will help to improve the international reputation of Australian wines and increase global competitiveness. Understanding of the important regional chemical parameters would allow grape growers and winemakers to optimise their viticultural and winemaking practices to preserve these characteristics of their terroir. Moreover, verifying the content in the bottle according to the label descriptions with a rapid method, has the potential to verify product provenance and counteract fraud in cases where wine of inferior/questionable quality or contaminated wine is presented as originating from Australia.

Keywords: Geographical origin, chemometrics, modelling, excitation-emission matrix

Introduction

Place of origin is a significant component in wine quality and a distinctive characteristic of a wine's sensory profile. Wine provenance and its embodiment of terroir is considered an important driver for consumer purchasing decisions (Warman and Lewis, 2019). In addition, wine is a luxury product that gains value from its terroir, so authentication of the geographical origin of wine has increasingly become a necessity in the wine industry to counter fraudulent activity. In Australia like other wine growing regions, the notion of geographical indication (GI) has been developed as ''one that identifies the wine as originating in a region or locality where a given quality, reputation or other characteristics of the wine is essentially attributable to the geographical origin''(Wine Australia, 2018). This allows differentiation of wines produced from winegrapes that are grown in different regions. However, when considering the geographical origin of wine, it is not simply the place where the winegrapes are grown that translates into its unique regionality. The location is underpinned by influences on grape cultivar from climate, soil, topography, viticultural practices etc, which relate to the broad concept of terroir (van Leeuwen and Seguin, 2006).

For geographical authentication, influence of terroir on regional variations can be described according to the differences of chemical components in wine (Roullier-Gall, *et al.,* 2014). These markers of geographical origin define the "identity" of a wine. Chemical measures such as elemental composition, stable isotope ratios, amino acid profile, grape and wine volatile compounds and polyphenols have been explored using range of advanced analytical methods to verify important regional parameters encompassed within the chemical composition of wines of provenance (Ranaweera, *et al.,* 2020a). However, in terms of practical application, it is necessary to consider a method that is rapid, accessible for in situ analyses, simple to implement, relatively low cost, and has high sensitivity and specificity. More recently, fluorescence spectroscopy has been explored as a tool to fulfil these requirements. This works by producing excitation emission matrices (EEMs) that provide a unique molecular fingerprint of each of the wine samples. Given the complexity of the datasets, multivariate data analysis methods (i.e., chemometrics) are often utilised to identify the patterns or classification groups of a particular wine. This method was successfully applied for the geographical authentication of a set commercial Cabernet Sauvignon wines (Ranaweera *et al.,* 2020b).

In the current study, we aimed to apply fluorescence spectroscopy in combination with various multivariate algorithms to develop a robust authentication model for Australian Shiraz wines produced at a commercial scale from different South Australian regions. This was undertaken using absorbance-transmittance with EEM (known as the A-TEEM technique) to further assess the effectiveness of this tool for regional authentication.

Materials and Methods

A total of 134 samples of unfinished (2019 vintage) commercial Shiraz wines from six different GIs of South Australia (Barossa Valley, Clare Valley, Eden Valley, Langhorne Creek, McLaren Vale, and Riverland) were analysed in duplicate by the A-TEEM technique according to Ranaweera *et al.* (2020b). An Aqualog spectrophotometer (Horiba Scientific, Version 4.2) was used to record the absorbance spectra, hue, intensity CIE L*a*b, CIE 1931, and EEMs of the samples. In the data acquisition process of the Aqualog, all EEMs were pre-processed prior to statistical analysis by normalising according to the water Raman scattering units for the specified emission conditions and correcting for the influence of inner filter effects (IFE) and Rayleigh masking. Multivariate algorithms, including partial least square discriminant analysis (PLSDA), support vector machine discriminant analysis (SVMDA), and extreme gradient boost discriminant analysis (XGBDA) were examined for the classification of wines. The data were pre-processed with different options including mean centreing, autoscaling, and generalised least squares weighting. Data were then compressed by PCA or PLS regression, applying the pre-processing method that provided the highest classification probability for each assigned class. The effectiveness of the cross-validated modelling techniques (Venetian blinds method; k=10) was compared by considering the accuracy of the predictions. Data analysis was undertaken using Solo software (version 8.8.1, Eigenvector Research, Inc., Manson, WA, USA).

Results and Discussion

EEM contour maps of the Shiraz wine samples were obtained from all the different regions, similar to the example shown for Barossa Valley and Eden Valley (Figure 1). The EEM signals arise from the fluorophores present in the wine such as phenolic compounds. Although subtle, definite differences can be seen in the each EEM maps (hence the notion of these being a molecular fingerprint), especially around excitation/emission wavelengths (EX/EM) of 275/320 nm. However, most components of wine have broad overlapping fluorescence excitation and emission spectra in the UV and visible range (Gilmore *et al.,* 2017), therefore it is necessary to employ multivariate statistical analysis to extract the information and apply it for classification according to origin.

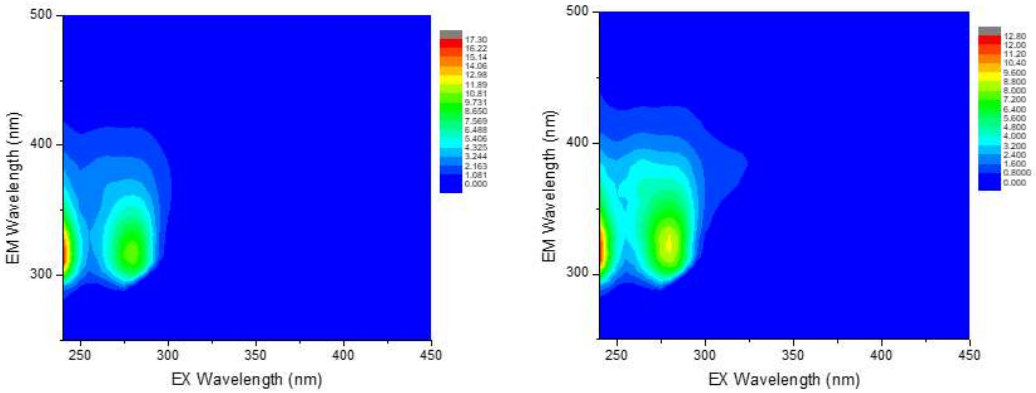


Figure 1. Example of EEM contour maps of Shiraz wine from Barossa Valley (left) and Eden Valley (right).

Multivariate data analysis with different supervised machine learning algorithms was undertaken to explore the potential for assigning samples to the correct class (i.e., region) according to the fluorescence measurements. PLSDA is a linear classification method recognised as a useful feature selector and classifier in food authentication (Song *et al.,* 2018). On the other hand, SVMDA is also a well-known learning algorithm, which represents a nonlinear classification technique (Song *et al.,* 2018). In comparison, XGBDA has yet to be applied broadly in a biological setting, but has previously been able to classify commercial Cabernet Sauvignon wines from three different regions of Australia and Bordeaux with 100% accuracy using fluorescence data (Ranaweera *et al.,* 2020b). These modelling techniques were applied to the Shiraz wines in the present study, yielding classification results as summarised in Table 1.

Table 1. Confusion matrix showing the cross-validation results of XGBDA, PLSDA and SVMDA models of EEMs for the different wine regions.

	Actual Class					
	Barossa Valley	Clare Valley	Eden Valley	Langhorne Creek	McLaren Vale	Riverland
Predicted Class						
			XGBDA			
Barossa Valley	56	0	0	0	0	0
Clare Valley	0	16	0	0	0	0
Eden Valley	0	0	22	0	0	0
Langhorne Creek	0	0	0	30	0	0
McLaren Vale	0	0	0	0	26	0
Riverland	0	0	0	0	0	118
Accuracy %	**100**	**100**	**100**	**100**	**100**	**100**
			PLSDA			
Barossa Valley	56	0	4	0	0	0
Clare Valley	0	10	0	0	0	0
Eden Valley	0	0	18	0	0	0
Langhorne Creek	0	0	0	29	2	0
McLaren Vale	0	6	0	1	24	0
Riverland	0	0	0	0	0	118
Accuracy %	**100**	**63**	**82**	**97**	**92**	**100**
			SVMDA			
Barossa Valley	54	0	6	1	1	0
Clare Valley	0	13	0	1	2	0
Eden Valley	2	0	16	0	0	0
Langhorne Creek	0	0	0	27	2	0
McLaren Vale	0	3	0	1	21	0
Riverland	0	0	0	0	0	118
Accuracy %	**96**	**81**	**73**	**90**	**80**	**100**

XGBDA was by far the best performing model, with cross-validation affording 100% correct classification of Shiraz wines from all of the tested regions (Table 1). This is in accordance with the previous study of Cabernet Sauvignon wines (Ranaweera *et al.,* 2020b). On the other hand, PLSDA showed 100% correct classification for Barossa Valley and Riverland samples and 97% accuracy for Langhorne Creek samples, with only one misclassified sample as McLaren Vale (Table 1). However, Eden Valley and Clare Valley samples were among the lowest in accuracy for PLSDA (82% and 63%). With SVMDA, similarly to other two methods, Riverland showed 100% correct classification and 96% for Barossa Valley. Yet there was relatively poor performance for Eden Valley samples, which showed the lowest accuracy using SVMDA (73%), and were misclassified as Barossa Valley. Langhorne Creek, McLaren Vale, Clare Valley gave an accuracy of 90%, 81% and 80%, respectively, with SVMDA.

These results in combination with colour measures from A-TEEM will lead to further investigation of chemical drivers behind this classification in future studies. Overall, the outcomes highlighted the integral capability of XGBDA for effective classification without overfitting the data and for parallel processing of unbalanced datasets, as reported previously (Ranaweera *et al.,* 2020b). Other model performance parameters, including sensitivity, specificity, precision, and F1 score, were also considered (data not shown), with XGBDA showing the highest values (1.0) for all of the parameters for each of the regions compared to PLSDA and SVMDA (< 1.0) and hence verifying this approach as the best performing classification model.

Conclusion

The results emphasised that the A-TEEM technique, in combination with the powerful multivariate tool XGBDA, can be highly effective in the authentication of wines according to their geographical origin. Moreover, the additional data obtained from A-TEEM can provide useful information on the typical colour and phenolic measures undertaken for red wine. Ultimately, unveiling inter-regional variations could be applied in the future to understand the influence of terroir for Australian wine regions. This will be beneficial to the optimisation or preservation of regional expression in wine and to improve the economic value of wines arising from different regions.

Acknowledgments

We are hugely thankful to Warren Birchmore, Viticulture Systems Manager from Accolade Wines, for the assistance with the sample collection from the different wineries owned by Accolade Wines. We acknowledge the financial support provided by an Australian Government Scholarship and Wine Australia supplementary scholarship (WA Ph1909).

References

Gilmore, A., Akaji, S., Csatorday, K., 2017. Spectroscopic analysis of red wines with A-TEEM molecular fingerprinting. Readout, 2: 41-48.

Ranaweera, RKR., Souza Gonzaga, L., Capone, DL., Bastian, SEP., Jeffery, DW., 2020a. Authenticity andtraceability in the wine industry: From analytical chemistry to consumer perceptions. In: Smithers, G. (Ed.), *Reference Module in Food Science*. Elsevier: United Kingdom.

Ranaweera, RKR., Gilmore, AM., Capone, DL., Bastian, SEP., Jeffery, DW., 2020b. Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling. Food Chemistry, 335: 127592.

Roullier-Gall, C., Lucio, M., Noret, L., Schmitt-Kopplin, P., Gougeon, RD., 2014. How subtle is the "terroir" effect? Chemistry-related signatures of two "Climats de Bourgogne". PLoS One, 9: 97615.

Song, W., Wang, H., Maguire, P., Nibouche, O., 2018. Collaborative representation based classifier with partial least squares regression for the classification of spectral data. Chemometrics and Intelligent Laboratory Systems, 182: 79-86.

van Leeuwen, C., Seguin, G., 2006. The concept of terroir in viticulture. Journal of Wine Research, 17: 1-10.

Warman, RD., Lewis, GK., 2019. Wine place research: Getting value from terroir and provenance in premium wine value chain interventions. International Journal of Wine Business Research, 31: 493-508.

Wine Australia, 2018. *Wine Australia Act 2013.* Australia.