R.K.R. Ranaweera, A. M. Gilmore, D.L. Capone, c, S.E.P. Bastian, D.W Jeffery

Application of fluorescence spectroscopy with multivariate analysis for authentication of Shiraz wines from different regions

Proceedings of the XIIIth International Terroir Congress, as published in IVES Conference Series, 2020, pp.1-5

XIIIth International Terroir Congress
17–18 November 2020
Virtual Congress | Adelaide, Australia

APPLICATION OF FLUOROSCENCE 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.

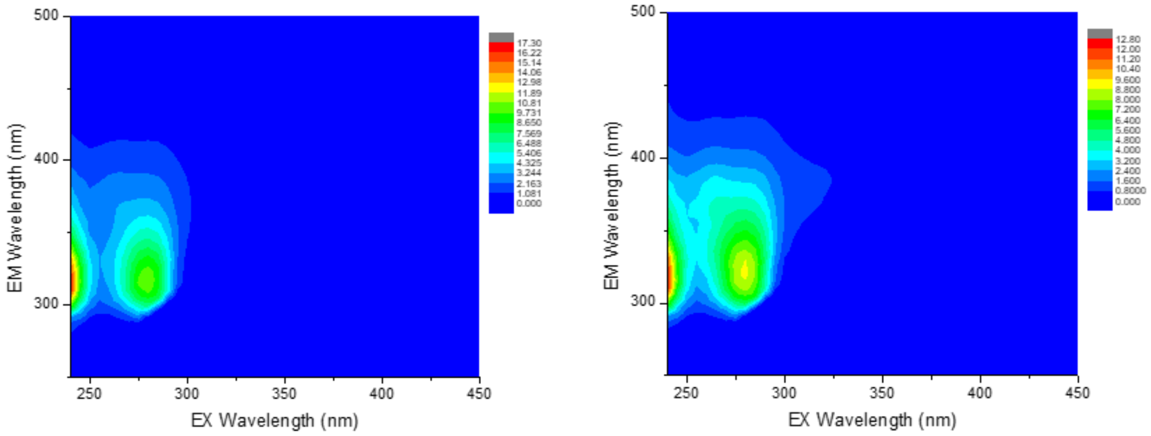

Figure 1. Example of EEM contour maps of Shiraz wine from Barossa Valley (left) and Eden Valley (right).

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.