Characterisation of grapevine berry samples with infrared spectroscopy methods and multivariate data analyses tools

by

Davirai M Musingarabwi

Thesis presented in partial fulfilment of the requirements for the degree of
Master of Science

at
Stellenbosch University
Institute for Wine Biotechnology, Faculty of AgriSciences

Supervisor: Prof Melané Vivier
Co-supervisor: Dr Hélène Nieuwoudt

March 2015

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 9 January 2015

Summary

Grape quality is linked to the organoleptic properties of grapes, raisins and wine. Many advances have been made in understanding the grape components that are important in the quality of wines and other grape products. A better understanding of the compositional content of grapes entails knowing when and how the various components accumulate in the berry. Therefore, an appreciation of grape berry development is vitally important towards the understanding of how vineyard practices can be used to improve the quality of grapes and eventually, wines.

The more established methods for grape berry quality assessment are based on gravimetric methods such as colorimetry, fluorescence and chromatography. These conventional methods are accurate at targeting particular components, but are typically multi-step, destructive, expensive, polluting procedures that might be technically challenging.

Very often grape berries are evaluated for quality (only) at harvest. This remains a necessary exercise as it helps viticulturists and oenologists to estimate some targeted metabolite profiles that are known to greatly influence chemical and sensory profiles of wines. However, a more objective measurement of predicting grape berry quality would involve evaluation of the grapes throughout the entire development and maturation cycle right from the early fruit to the ripe fruit. To achieve this objective, the modern grape and wine industry needs rapid, reliable, simpler and cost effective methods to profile berry development. By the turn of the last millennium, developments in infrared instrumentation such as Fourier-transform infrared (FT NIR) and attenuated total reflectance Fourier-transform infrared spectroscopy (ATR FT-IR) in combination with chemometrics resulted in the development of rapid methods for evaluating the internal and external characteristics of fresh fruit, including grapes. The advancement and application of these rapid techniques to fingerprint grape compositional traits would be useful in monitoring grape berry quality.

In this project an evaluation of grape berry development was investigated in a South African vineyard setting. To achieve this goal, Sauvignon blanc grape berry samples were collected and characterised at five defined stages of development: green, pre-véraison, véraison, post-véraison and ripe. Metabolically inactivated (frozen in liquid nitrogen and stored at -80^{o}C) and fresh berries were analysed with FT-IR spectroscopy in the near infrared (NIR) and mid-infrared (MIR) ranges to provide spectral data. The spectral data were used to provide qualitative

(developmental stage) and quantitative (metabolite concentration of key primary metabolites) information of the berries.

High performance liquid chromatography (HPLC) was used to separate and quantify glucose, fructose, tartaric acid, malic acid and succinic acid which provided the reference data needed for quantitative analysis of the spectra. Unsupervised and supervised multivariate analyses were sequentially performed on various data blocks obtained by spectroscopy to construct qualitative and quantitative models that were used to characterise the berries. Successful treatment of data by principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA) gave statistically significant chemometric models that discriminated the berries according to their stages of development. The loadings from MIR models highlighted the important discriminant variables responsible for the observed developmental stage classification. The best calibration models to predict metabolite concentrations were obtained from MIR spectra for glucose, fructose, tartaric acid and malic acid. The results showed that both NIR and MIR spectra in combination with multivariate analysis could be reliably used to evaluate Sauvignon blanc grape berry quality throughout the fruit's development cycle. Moreover, the methods used were fast and required minimal sample processing and no metabolite extractions with organic solvent. In addition, the individual major sugar and organic acids were accurately predicted at the five stages under investigation. This study provides further proof that IR technologies are robust and suitable to explore high-throughput and in-field application of grape compound profiling.

Opsomming

Druifkwaliteit word gekoppel aan die organoleptiese eienskappe van druiwe, rosyntjies en wyn. Baie vooruitgang is reeds gemaak in die begrip van druifkomponente wat belangrik is vir die kwaliteit van wyn en ander druifprodukte. ’n Beter begrip van die samestellende inhoud van druiwe behels om te weet wanneer en hoe die verskeie komponente in die korrel opgaar. ’n Evaluasie van druiwekorrel-ontwikkeling is dus uiters belangrik vir ’n begrip van hoe wingerdpraktyke gebruik kan word om die kwaliteit van druiwe, en uiteindelik van wyne, te verbeter.

Die meer gevestigde maniere vir die assessering van druiwekorrelkwaliteit is gebaseer op gravimetriese metodes soos kolorimetrie, fluoressensie en chromatografie. Hierdie konvensionele metodes is akkuraat om spesifieke komponente te teiken, maar behels tipies veelvuldige stappe en is prosesse wat destruktief en duur is, besoedeling veroorsaak, asook moontlik tegnies uitdagend is.

In baie gevalle word druiwekorrels (eers) tydens oes vir kwaliteit geëvalueer. Hierdie is steeds ’n noodsaaklike oefening omdat dit wingerdkundiges en wynkundiges help om die metabolietprofiele wat daarvoor bekend is om ’n groot invloed op die chemiese en sensoriese profiele van wyn te hê en dus geteiken word, te skat. ’n Meer objektiewe meting om druiwekorrelkwaliteit te voorspel, sou die evaluering van die druiwe dwarsdeur hulle ontwikkeling- en rypwordingsiklus behels, vanaf die vroeë vrugte tot die ryp vrugte. Om hierdie doelwit te behaal, benodig die moderne druiwe- en wynbedryf vinnige, betroubare, eenvoudiger en kostedoeltreffende metodes om ’n profiel saam te stel van korrelontwikkeling. Aan die einde van die vorige millennium het ontwikkelings in infrarooi instrumentering soos Fourier-transform infrarooi (FT NIR) en *attenuated total reflectance* Fourier-transform infrarooi spektroskopie (ATR FT-IR) in kombinasie met chemometrika gelei tot die ontwikkeling van vinnige metodes om die interne en eksterne kenmerke van vars vrugte, insluitend druiwe, te meet. Die vooruitgang en toepassing van hierdie vinnige tegnieke om ‘vingerafdrukke’ te bekom van die samestellende kenmerke sal nuttig wees vir die verbetering van druiwekorrelkwaliteit.

In hierdie projek is ’n evaluering van druiwekorrelontwikkeling in ’n Suid-Afrikaanse wingerdligging ondersoek. Ten einde hierdie doel te bereik, is Sauvignon blanc druiwekorrelmonsters op vyf gedefinieerde stadiums van ontwikkeling versamel en gekarakteriseer: groen, voor deurslaan, deurslaan, ná deurslaan en ryp. Metabolies geïnaktiveerde (bevrore in vloeibare stikstof en gestoor teen -80^{o}C) en vars korrels is met FT-IR spektroskopie in die naby infrarooi (NIR) and mid-infrarooi (MIR) grense

geanaliseer om spektrale data te verskaf. Die spektrale data is gebruik om kwalitatiewe (ontwikkelingstadium) en kwantitatiewe (metabolietkonsentrasie van belangrikste primêre metaboliete) inligting van die korrels te verskaf.

High performance liquid chromatography (HPLC) is gebruik om glukose, fruktose, wynsteensuur, appelsuur en suksiensuur te skei en te kwantifiseer, wat die verwysingsdata verskaf het wat vir die kwantitatiewe analise van die spektra benodig word. Ongekontroleerde en gekontroleerde meervariantanalises is opeenvolgend op verskeie datablokke uitgevoer wat met spektroskopie verkry is om kwalitatiewe en kwantitatiewe modelle te verkry wat gebruik is om die korrels te karakteriseer. Suksesvolle behandeling van die data deur hoofkomponent analise (*principal component analysis* (PCA)) en ortogonale parsiële kleinstekwadraat diskriminant analise (*partial least squares discriminant analysis* (OPLS-DA)) het statisties betekenisvolle chemometriese modelle verskaf wat die korrels op grond van hulle ontwikkelingstadia onderskei het. Die ladings vanaf die MIR-modelle het die belangrike diskriminantveranderlikes beklemtoon wat vir die klassifikasie van die waargenome ontwikkelingstadium verantwoordelik is. Die beste kalibrasiemodelle om metabolietkonsentrasies te verkry, is vanuit die MIR-spektra vir glukose, fruktose, wynsteensuur en appelsuur bekom. Die resultate toon dat beide die NIR- en MIR-spektra, in kombinasie met meervariantanalise, betroubaar gebruik kan word om Sauvignon blanc druiwekorrelkwaliteit dwarsdeur die vrug se ontwikkelingsiklus te evalueer. Verder is die metodes wat gebruik word, vinnig en het hulle minimale monsterprosessering en geen metabolietekstraksies met organiese oplosmiddel benodig nie. Daarbenewens is die vernaamste suiker en organiese sure individueel akkuraat voorspel op die vyf stadia wat ondersoek is. Hierdie studie verskaf verdere bewys dat IR-tegnologieë robuus en geskik is om hoë-deurset en in-veld toepassings van profielsamestelling van druiweverbindings te ondersoek.

This thesis is dedicated to Victoria and Winsley

Biographical sketch

Davirai was born in Masvingo (formerly Fort Victoria), Zimbabwe where he did undergraduate studies at the University of Zimbabwe. After university, a stint in the food and chemicals industry was followed by a Master of Science degree in Organic Chemistry from the University of Vermont, USA. He is a keen sporting enthusiast and loves watching soccer, tennis and cricket.

Acknowledgements

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- My advisor, Prof Melanè Vivier;
- My co-advisor, Dr Hélène Nieuwoudt;
- Dr Philip Young for providing expertise regarding the vineyard layout, who performed sampling from the field and contributed to processing of the frozen berry samples; also for contributing to the editing of text and diagrams in the research paper that was submitted, based on the research results presented in Chapter 3;
- Dr Liezel Gouws is acknowledged for organic extraction of grape berry metabolites for the reference HPLC data presented in Chapter 3;
- Dr Hans Eyéghè-Bickong is thanked for performing the HPLC analysis of the reference samples presented in Chapter 3;
- Ms Varsha Premagasar for her technical support in milling and storage of frozen berry samples;
- Prof Martin Kidd, for assistance with statistical analysis and interpretations;
- Prof Johan Trygg, for immense assistance with chemometrics;
- The IWBT (Stellenbosch University) and Winetech (South Africa), for financial support

Preface

This thesis is presented as a compilation of four chapters. Each chapter is introduced separately and is written according to the style of the Food Chemistry journal. The student was responsible for all the experimental work, unless otherwise stated (see below for Chapter 3), drafting and finalising the thesis, with study guidance with regards to conceiving of the study, experimental design, interpretation of the data and appropriate data presentation, as well as editing of the thesis provided by the two supervisors.

Chapter 3 was submitted as a journal article in Food Chemistry and the following authors are listed: Davirai M. Musingarabwi, Hélène H. Nieuwoudt, Philip R. Young, Hans A. Eyéghè-Bickong, and Melanè A. Vivier. For this multi-author paper, DMM conducted all experimental work regarding spectroscopy analysis and subsequent multivariate data-analysis and drafting of the paper; HHN supervised the details of the spectroscopy and chemometrics analyses and contributed to data-interpretation and the editing and final draft of the paper; PRY contributed samples and to data analysis, as well as preparing the final drafts of the submission; HEB performed HPLC analysis on the reference samples and MAV conceived the study, supervised the student and edited the drafts of the paper.

Chapter 1	General introduction and project aims
Chapter 2	Literature review: A review of IR spectroscopy and chemometrics in their application to grapevine research
Chapter 3	Research results: A rapid qualitative and quantitative evaluation of grape berries at various stages of development using Fourier-transform infrared spectroscopy and multivariate data analysis
Chapter 4	General discussion and conclusions, future perspectives and evaluation of the study

Table of Contents

CHAPTER 1

Introduction and project aims

1.1 Introduction

The grape berry is of great significance to humans as a fruit product and as the chief ingredient in wine production. Consequently, its development and maturity has received and continues to receive considerable scientific scrutiny (Conde et al., 2007). In-depth grapevine research is also aimed at improving grape vintage quality (Falcão et al., 2008). This includes the study of targeted grape metabolites that are considered fundamental to berry development and perceived as important to improve vineyard management and production of optimally ripe grapes (Ben-Ghozlen et al., 2010).

In premium wine production, the quality of the vintage quality is rated higher than its yield where a high quality vintage is associated with as near uniform grape development as possible (Keller et al., 2006). Some of the factors that influence vintage quality include initial site selection, viticultural practices (irrigation, crop density, canopy management and staggered harvest) and the particular season (Jackson & Lombard, 1993). Concerted efforts to produce an optimal crop is always subject to the heterogeneity in the berries within a bunch, berries within a vine and consequently within the vineyard (Dal Santo et al., 2013). Thus, in practice it is difficult to determine when a vineyard with unknown variation in berry development is at its best possible ripeness.

Although grape quality at harvest is the main factor for conditioning wine quality, grape maturation is a continuous process, which goes through various development stages and impacts on the quality of the fruit at harvest (Meléndez et. al., 2013). Therefore, in order to produce more desirable grapes, practitioners in the grape and wine industry would need to assess grape quality throughout berry development so that they can design how the various viticultural strategies would work best in their own environments. To this end they would need access to simpler yet effective techniques that can help in fruit assessment.

Traditionally, grape berries were evaluated for sugars and organic acids using conventional gravimetric methods like chromatography (Hunter et al, 1991), colorimetry (Celotti & De Prati, 2005) or fluorescence (Tuccio et al., 2011). These methods are and are likely to remain the standard bearers for in-depth grapevine research, though they lag behind the demands of modern requirements of the

grape and wine industry (Herrera et al., 2003). A major challenge for today's oenologist is the quantification of wine grape composition prior to vinification (Herrera et al., 2003). Due to time constraints at harvest, such quantification cannot be determined by conventional techniques as those previously mentioned. Much faster methods are needed for this. To this end, comparatively faster, non-selective methods like nuclear magnetic resonance spectroscopy have been used as alternatives to evaluate grape berry quality (Ali et al., 2011). But it is the recent developments in IR spectroscopy and multivariate data analysis that have provided momentum towards achieving the requirements for even faster analysis of grape berries (Gishen, et al., 2005).

By the turn of the last millennium there was a growing global trend to analyse a food sample in its entirety and not just targeted characteristics so, food quality measurement techniques have been carried with the aid of multivariate based measurements of chemical and physical properties (Martens & Naes, 1989; Jaumot et al., 2004). Spectrometric techniques in the NIR and MIR regions have been used to evaluate IR active compounds in many sample matrices as they offer some advantages like increased speed of analysis, cleaner processing and simplicity as compared to the conventional means of analysis (Bauer et al, 2008). In this project, NIR and MIR spectroscopy and multivariate analyses were used to evaluate grapes in both qualitative (development stage) and quantitative terms (targeted metabolite concentration at the different developmental stages).

1.2 Project background and specific objectives

Prior to the onset of this project, a chromatographic method was optimised in our laboratory as part of an integrated approach towards tracking berry development and ripening using targeted metabolites tartaric acid, malic acid, succinic acid, fructose and glucose (Eyéghè-Bickong et al., 2012). The method achieved a major reduction in sample size (from the gram level to the mg level) and managed to track grape berry development and ripening at different stages of berry growth. It was evaluated and found to be suitable for targeted profiling grapevine berry metabolites during the different berry developmental stages. Due to the work-flow associated with the method, grapes had to be sampled, flash-frozen, milled and

kept frozen (- 80^{o}C) till the analysis could be carried out. A spectroscopic method was proposed to complement the chromatographic approach to analyse the berries faster and probe if it could be extended to fresh fruit.

The project was established to design workable and reproducible NIR and MIR spectroscopic workflows that could be used for evaluating grape berry development and maturation using faster, cost effective and greener methods. Initially, the focus was to evaluate grape berry development using flash frozen samples and if satisfactory results were obtained, the method further assessed to determine if the same techniques used for frozen berries could be extended to fresh grapes.

Both FT-NIR and ATR FT-MIR spectroscopy were used in this project. A look into the scientific literature showed that FT-NIR was evidently the predominant spectroscopic technique in the food industry. The project interests included the simultaneous determination of targeted metabolites with similar molecular structures like glucose and fructose, tartaric and malic acid present in the grape berry matrix. NIR spectroscopy is an established non-destructive technique with numerous applications in the agricultural, horticultural and food sectors (Quilitzsch et al., 2005; Gishen et al., 2005). FT-NIR molecular vibrations is characterised by overtone and combination bands for functional groups such as C-H, N-H and O-H whilst ATR FT-MIR is more sensitive to a wider range of functional groups including C-O, C-N, N-H, O-H and C-S groups not detectable in the NIR range (McClure 2003; Davis & Mauer, 2010). The sensitivity of ATR FT-MIR spectroscopy has been tested with complex matrices like dairy fat and its accuracy has seen ATR FT-MIR as the first (if not the only) spectroscopic analytical technique to be approved as an official American Oil Chemists' Society method (AOCS, 1999). Therefore the evaluation of both FT-NIR and ATR FT-MIR methods for their ability to produce different metabolite signals in different concentration combinations at different grape berry stages was deemed to bring in added scope to our study .

This thesis was based on studies where the biological phenomenon of grape berry growth and development was evaluated using IR spectroscopy, HPLC and multivariate analysis. The specific aims and associated tasks of the study were as follows:

1. To employ a rapid and powerful analytical workflow that can be used to study grape berry development at five distinct stages.

2. To present multivariate analysis tools and strategies that can be used to evaluate the spectral and chromatographic data.

3. To construct multivariate models that can simplify the spectral data:
    a. qualitative analysis: unsupervised multivariate analysis (principal component analysis, PCA);
    b. qualitative analysis: supervised multivariate analysis (orthogonal partial least squares discriminant analysis , OPLS-DA);
    c. quantitative analysis: partial least squares (PLS) models.

4. To come out with data outputs that can be used to evaluate the spectral and chromatographic information:
    a. classification of grape berry according to developmental stage (qualitative models);
    b. prediction of grape berry metabolites (quantitative models) and comparison with reference data regarding these key metabolites.

1.3 Cited Literature

Ali, K., Maltese, F., Fortes, A.M., Pais, M.S., Choi, Y.H., & Verpoorte, R. (2011). Monitoring biochemical changes during grape berry development in Portuguese cultivars by NMR spectroscopy. *Food Chemistry, 124,* 1760-1769.

American Oil Chemists' Society (AOCS) (1999). *Official Methods and Recommended Practices, 5th ed.,* Firestone, D. (Ed), Champaign: AOCS (Official method Cd 14d-99).

Bauer, R., Nieuwoudt, H., Bauer, F.F., Kossmann, K., Koch, K.R., & Esbensen, K.H. (2008). FTIR spectroscopy for grape and wine analysis. *Analytical Chemistry, 1,* 1371-1378.

Ben-Ghozlen, N., Moise, N., Latouche, G., Martino, V., Mercier, L., Besançon, E., & Cerovic, Z.G. (2010). Assessment of grapevine maturity using a.new portable sensor: non-destructive quantification of anthocyanins. Journal *International des Sciences de la Vigne et du Vin (special issue MacroWine),* 1-8.

Celotti, E., & De Prati, G.C. (2005). The phenolic quality of red grapes at delivery: Objective evaluation with colour measurements. *South African Journal for Enology and Viticulture, 26,* 75-82.

Conde, C., Silva, P., Fontes, N., Dias, A.C.P., Tavares, R.M., Sousa, M.J., Agasse, A., Delrot, S., & Gerós, H. 2007. Biochemical changes throughout grape berry development and fruit and wine quality*. Food.* 1-18.

Dal Santo, S., Tornielli, G.B., Zenoni, S., Fasoli, F., Farina, L., Anesi, A. Guzzo, F., Massimo Delledonne, M., & Pezzotti, M. (2013). The plasticity of the grapevine berry transcriptome*. Genome Biology, 14*, 1-18.

Eyéghè-Bickong, H.A.; Alexandersson, E.O., Gouws, L.M., Young, P.R., & Vivier, M.A. (2012). Optimization of an HPLC method for the simultaneous quantification of the major sugars and organic acids in grapevine berries*. Journal of Chromatography B, 885-886,* 43-49.

Falcão, L.D., Chaves, E.S., Burin,V.M., Falcão,A.P., Gris,E.F., Bonin, V., & Bordignon-Luiz, M.T. 2008. (2008). Maturity of Cabernet Sauvignon berries from grapevines grown with two different training systems in a new grape growing region in Brazil*. Ciencia e Investigación Agraria.* 35, 271-282.

Gishen, M., Dambergs, R.G., & Cozzolino, D. (2005). Enhancing spectroscopy with chemometrics. *Australian Journal of Grape and Wine Research 11,* 296-305.

Herrera, J., Guesalaga, A., & Agosin, E. (2003). Shortwave near infrared spectroscopy for non-destructive determination of maturity of wine grapes. *Measurement Science and Technology, 14,* 689-697.

Hunter J.J., Visser, J.H., & De Villiers, O.T. (1991). Preparation of grapes and extraction of sugars and organic acids for determination by high performance liquid chromatography. *American Journal of Enology and Viticulture, 43,* 237-244.

Jackson, D.I., & Lombard, P.B. (1993). Environmental and management practices affecting grape composition and wine quality - a review. *American Journal of Enology and Viticulture, 44,* 409-430.

Jaumot, J., Vives, M., & Gargallo, R. (2004). Application of multivariate resolution methods to the study of biochemical and biophysical processes. *Analytical Biochemistry 327,* 1-13.

Keller, M., Smith, J.P., & Bondada, B.R. (2006). Ripening grape berries remain hydraulically connected to the shoot. *Journal of Experimental Botany,* 1-11.

Martens, H., & Naes, T. (1989). *Multivariate Calibration.* Toronto: John Wiley & Sons (Chapter 1).

McClure, W.F. (2003). 204 years of near infrared technology: 1800–2003. *Journal of Near Infrared Spectroscopy 11,* 487-518.

Meléndez, E., Ortiz, M.C., Sarabia, L.A., Íñiguez, M., & Puras, P. (2013). Modelling phenolic and technological maturities of grapes by means of the multivariate relation between organoleptic and physicochemical properties. *Analytica Chimica Acta, 761,* 53-61.

Quilitzsch,R., Baranska, M., Schulz, H., & Hoberg, E (2005). Fast determination of carrot quality by spectroscopy methods in the UV-VIS,NIR and IR range. *Journal of Applied Botany and Food Quality 79,* 163-167.

Tuccio, L., Remorini, D., Pinelli, P., Fierini, E., Tonutti, P., Scalabrelli, G., & Agati, G. (2011). Non-destructive detection of grape anthocyanins. *Australian Journal of Grape and Wine Research 17,* 181-189.

CHAPTER 2

Literature review: A review of infrared spectroscopy and chemometrics in their application to grapevine research

2.1 Introduction

Grapes, members of the genus *Vitis* of the family Vitaceae, were among the earliest fruit grown by man and to date, are the most economically important fruit crop in the world with *Vitis vinifera* and its hybrids accounting for over 90% of the world grapes (Soyer et al., 2003). Global grape production figures exceeded 75 million tonnes by 2012 (OIV, 2013). Of the total grape harvest, 70% is used for making wine, 27% is for direct consumption as table grapes, 2% is used for making raisins and less than 1% is used for making musts or distillates (Pavloušek & Kumšta, 2011). In addition to their economic importance, an increasing number of health benefits *e.g.* antioxidant, cardio-protective, anti-inflammatory and anti-cancer activities have been attributed to grapes and its downstream products (Ali et al., 2011). The assessment of fruit composition is a cornerstone of fruit management and the processes of fruit development and maturation could be used to project crop quality patterns (Butz et al., 2005). With the bulk of grape berries set for wine production, it is fundamental that targeted grape compositional analysis is carried out during cultivation if the desired wines are to be produced (Arana et al., 2005).

Fruits, including grapes, display dynamic qualitative (*e.g.* colour, size) and quantitative changes (synthesis or accumulation and degradation of metabolites) during the process of development and ripening (White, 2002). In grapes, these changes are helpful to the viticulturist, the vintner and the researcher as they can be used by the producers to assess their preferred fruit quality throughout the growing season. Currently viticulturists are concerned by berry size, acidity, colour, volatile and non-volatile contents of the grape berry as it grows and ripens (Ali et al., 2011). The grape berry has the capacity to accumulate a wide diversity of compounds including sugars, organic acids, phenolic, flavour and other compounds (Coombe & McCarthy, 2000). These compounds undergo some significant compositional changes as the fruit develops (Gholami et al., 1995).

Typically, berry development and ripening is characterised by asynchrony between the berries on a bunch, the bunches on a vine and the vines within a vineyard (Dal Santo et al., 2013). Fruit classification methods are sometimes implemented to monitor berry heterogeneity in attempts aimed at providing

possible trends in the accumulation and metabolism of major grape berry compounds (Garcia de Cortazar-Atauri et al., 2009). The existence of in-field grape variability has led to the development of new concepts for vineyard management and staggered harvests. However to be routinely implemented in grapevine production, the practical implementation of these concepts requires the use of simpler, reliable and cost effective technologies (Rolle et al., 2012).

Traditionally some of the grape berry quality markers such as sugars and organic acids were determined gravimetrically by methods such as chromatography, fluorescence and colorimetry amongst other techniques as these provide highly accurate and reliable information (Hunter et al., 1991; Ali et al., 2011; Celotti & De Prati, 2005; Tuccio et al., 2011). In addition to these techniques, non-selective approaches like nuclear magnetic resonance (NMR) spectroscopy are finding increasing application in grapevine research (Ali et al., 2011). The gravimetric methods remain the workhorses of more detailed grapevine research but they lag behind the demands of modern global grape and wine markets (Herrera et al., 2003). The growing global awareness by consumers for high-end fruit attributes like taste, nutritional content and product authenticity has prompted industry's demand for more innovative tools that can rapidly assess grape products for a number of qualitative and quantitative aspects (Butz et al., 2005). Modern viticulturists could be better served by simpler low-cost analytical techniques that are non-invasive or requiring minimum sample processing if they are to keep abreast with producer and consumer expectations (González-Caballero et al., 2010).

This review aims to provide a synopsis of some of the most relevant research topics related to the application of IR spectroscopy and chemometrics in grape berry technology. The review is divided into four sections (2.2 – 2.5) with each section looking at a particular theme. Section 2.2 gives an overview of the theory behind grape berry development processes in a manner that puts into context the desirability for newer technologies in this field. Section 2.3 gives a historical perspective of IR radiation and describes how this type of radiation has developed as an analytical technique in areas like grapevine research. A selection of important time-related developments in IR spectroscopy is provided: from the time IR was

first used as an analytical tool in crop sciences up to its more recent applications in among other scientific fields, grape berry technology. Section 2.4 discusses chemometrics, highlighting its principles and the tools for data analysis. Section 2.5 looks at how developments in chemometrics and IR spectroscopy have been combined to provide the powerful investigative platform they are today and gives some selected applications of IR spectroscopy and chemometrics in grapevine research. A number of examples constituting selected research endeavours in which IR spectroscopy and chemometrics were applied to achieve a variety of outcomes in the advancement of rapid grape berry analysis are considered. The list is not meant to be exhaustive but is presented to feature published works that cover some recent developments that highlight different versions of how IR spectroscopy and chemometrics were successfully used in grapevine research.

2.2 Grape berry development

A universal objective in viticulture over the years has been the desire to produce as uniformly ripe a crop as possible (Reynolds, 2010). The monitoring of grape berry development and eventual ripening is indispensable if reasonable measures of sustaining desired vineyard quality are to be implemented. Grape berries start off as small, hard and acidic berries with little sugar, proceeding through a series of developmental stages until when they are much larger, softer, less acidic, sweeter, strongly flavoured and coloured (Conde et al., 2007).

Over the years, a lot of scientific efforts were dedicated to gain a better understanding of the complexities that accompany the physical and biochemical changes occurring in grape berries as they develop and mature (Coombe, 1992). Like all fruits, grapes are seasonal and the assessment of their metabolite compositional changes during fruit growth would immensely benefit from the incorporation of quicker, reliable, cheaper, greener and timely protocols that can be used to expedite the monitoring of berry development, ripening and storage (Muñoz-Robredo et al., 2011).

Eventual grape organoleptic quality is greatly but not entirely dependent on the content and composition of the most abundant metabolites like sugars and organic

acids (Soyer et al., 2003). Fructose and glucose are the major sugars in the berry whilst tartaric and malic acids are the dominant organic acids in grapes, accounting for over 90% of the total acidity in grapes, the remainder is made up from minor organic acids including acetic, ascorbic, cinnamic, citric, isocitric, formic, fumaric, galacturonic, gallic, glutaric, glyceric, benzoic, α-ketoglutotaric, lactic, mandelic, oxalacetic, oxalic, phosphoric, pyrolidone carboxylic, pyruvic, salicylic, shicimic and succinic acids (Topalovic & Mikulic-Petkovsek, 2010). At any one time, the concentrations of these primary metabolites in the fruit give a significant indication of flavour quality and maturity (Coombe, 1992).

Chemical composition is one of the most important criteria used to define fruit quality (Soyer et al., 2003). Grape chemical composition is sensitive to among other things environmental and geographical factors, grape variety and vintage (Jackson & Lombard, 1993). This sensitivity demands that reliable spatially appropriate data on grape berry quality over the entire berry growth period be more readily available to viticulturists and wine producers (Teixeira et al., 2014). To this end, the continuous assessment of grape chemical composition is particularly useful to objectively assess grape berry development. Grape chemical composition has been well assessed by use of gravimetric methods that include chromatography (Hunter et al., 1991), colorimetry (Celotti & De Prati, 2005) and fluorescence (Tuccio et al., 2011). In order to monitor the accumulation/metabolism of grape metabolites as a season progresses, faster, simpler yet reliable and cost effective technologies would be important.

Grape berry development has been reported to consist of two successive sigmoidal cycles that are separated by a lag phase (Coombe, 1992). The first cycle involves the setting of the fruit and is ensued by rapid cell division, accumulation of organic acids, principally malic, tartaric and hydroxycinnamic acids; at this stage the berry is hard, green and slow in growing (Coombe, 1992). The second cycle starts at véraison and is characterised by exponential sugar accumulation, berry softening, colouring, and remarkable increase of berry size (over three-fold expansion), respiration of tartaric acid and malic acid and ends at engustment, at which point phenolics and flavour compounds are formed (Coombe, 1992). Other

descriptions of grape berry growth are replete in the literature but are invariably hinged to the double sigmoidal curves.

A widely accepted system to define grape phenology is the Eichhorn-Lorenz (E-L) system that is divided into 47 E-L numbers with each E-L number representing a growth stage (Coombe, 1995). Flowering is marked at E-L 19 and the actual berry development sets off at E-L 27 after which the berry goes through over a dozen stages culminating in ripening at E-L 38 (Coombe, 1995). Most grapevine research based on IR spectroscopy has been concentrated towards the maturation and ripening stages as these are the times that mainly affect harvesting and wine-making decisions.

Grapes display some asynchronous growth patterns at bunch vine and vineyard levels (Dai et al, 2013). This unsynchronised nature of grape berry development leads to in-field grape variability and has a significant impact on grape berry sampling design for example, and it is increasingly recognized that updated concepts, workflows and technologies are required to account for this complicating factor (Alexandersson et al., 2014).

2.3 Infrared spectroscopy

Infrared (IR) radiation (12500 – 40cm^{-1}, 800 - 250000 nm) is part of the electromagnetic spectrum lying between the visible and microwave regions and is defined as near infrared (NIR) (12500 – 4000 cm^{-1}, 800 - 2500 nm), mid-infrared (MIR) (4000 – 400 cm^{-1}, 2500-25000 nm) and far infrared (FIR) (400 – 40 cm^{-1}, 25000 - 250000 nm) (Simon, 1966; Davis & Mauer, 2010). In general, the boundaries between the near, mid and far-infrared regions are not agreed upon and they marginally vary between authors. The astronomer Sir William Herschel was first to recognize the existence of the IR region in 1800 (Walker, 2000). The discovery offered a potential new analytical tool to study organic and biochemical molecules though interest in IR radiation lay dormant for another century till William de Wiveleslie Abney and Edward Robert Festing photographed absorption spectra of 52 compounds and correlated the absorption bands of some organic groups in the molecules they were investigating (Dufour, 2009).

In 1903 William Weber Coblentz laid the groundwork for IR spectroscopy as it is known today when he investigated the spectra of hundreds of organic and inorganic substances and published a list of functional group frequencies (Barth & Haris, 2009). Analytical applications of IR spectroscopy started in 1949 with the USA Department of Agriculture pioneering a research project that evaluated egg quality (Duffour, 2009). The first quantitative study related to bio-organic compounds was published shortly after this pioneering work and the development of analytical methods based on NIR spectroscopy started around 1968 in the Karl Norris laboratory (Barth & Haris, 2009). Since then, low cost IR analytical instruments that are desirable for metabolite content evaluation in fruits, including the grape berry have been developed.

IR spectroscopy has several features that make it ideal for studying bioorganic molecules such as those found in grape berries. Some of its more attractive features include easy sample preparation; ability to obtain detailed information on the chemical structure of compounds at the molecular and functional group level; good reproducibility; and, in many cases, non-invasive approaches (Adato & Altug, 2013). The utility of IR spectroscopy in fruit quality evaluation has been enhanced by its combination with chemometrics (Butz et al., 2005). This combination has seen the techniques grow tremendously since the 1970s and at the same time attracting the attention of numerous research groups as an increasingly desirable method of choice in grapevine research (Butz et al., 2005; Varmuza & Filzmoser, 2009).

The NIR region is the first electromagnetic spectral region that exhibits absorption bands related to molecular vibrations and is characterised by harmonics and combination bands widely used for compositional analyses of food products (Dufour, 2009). NIR spectra are dominated by absorption bands associated with the functional groups C-H, O-H, or C-O and the sum of these vibrations generate combination bands or overtones which are approximate multiples of the fundamental frequencies (Williams & Norris, 2001). This makes NIR spectra far more complex than they appear. The MIR region is the main region of vibrational spectroscopy and is the region that contains most of the information that allows for bioorganic molecules to be identified together with the portrayal of the structures

and conformations of molecules such as proteins, saccharides, organic acids and lipids (Dufour, 2009). Both NIR and MIR spectroscopy are very useful for analysing high vibrating molecules like those in fluids whilst FIR is more useful for the analysis of low vibrating molecules especially inorganic compounds (Brügel, 1962).

Depending on the particular application and sample type, an IR spectrum can be measured in one of five major modes *viz* absorbance/transmittance, interactance, transflectance, diffuse transmittance and diffuse reflectance (Huang et al., 2008). In practice the most preferred IR analytical methods are absorption/transmission and reflectance (Brügel, 1962; Arzamastsev et al., 2008). Absorption/transmission IR spectroscopy is dependent on how strongly light is absorbed by the sample at each frequency whilst reflectance works by introducing IR light into a sample at an angle that exceeds the critical angle for internal reflection to create an evanescent wave at the reflecting surface and simultaneously triggers the recording of a spectrum (Arzamastsev et al., 2008).

When a sample is put through an IR instrument, a spectrum is produced as the molecular bond vibrates at the same frequency as that of the IR radiation and is recorded as a continuum consisting of a large number of primary data points that can be interpreted by chemometric tools (Blanco & Villarroya 2002). A typical fresh produce spectrum contains a number of peaks with several shoulders (Williams & Norris, 2001). The peaks are composites of numerous individual bands that cannot be resolved through visual inspection alone (Shenk et al, 2007).

Despite its distinct attractive features, IR spectroscopy has some notable drawbacks. The Lambert-Beer law (absorption is higher the larger the distance covered) limits the scope of IR spectroscopy as IR radiation can only penetrate finite distances of given material and the limit imposed by small absorption cross-sections makes IR signals weaker (Brügel, 1962). This renders IR spectroscopy regarded as unreliable to measure substances that occur at levels below 0.2 mg/g (Bauer et al., 2008). In aqueous measurements (prevalent in most fruit products), water presents a major obstacle for IR analysis as the water O-H bending absorption can mask most other sample signals to the level that they may appear non-existent (Adato & Altug, 2013).

Direct interpretation of IR spectra may be arduous especially in the NIR region, as some of the absorption peaks are emanating from complex overtone and high frequency combinations from primary absorptions (Kelly et al., 2004). However, the major chemometric softwares such as SIMCA (Umetrics, Sweden), Unscrambler (Camo, Norway) and MatLab (Mathworks, USA) have spectral filters that can remove undesirable variation emanating from instrumental imperfections, light scattering, or baseline shifts adding to the sample signal (Givens et al., 1997). Usually, NIR spectra would need the application of spectral filters before data processing due to the weaker signals emanating from overtones between 10000 - 5200 cm^{-1} and combinations between 5200 - 4000 cm^{-1} (Gishen et al., 2005). MIR spectra result from the fundamental stretching, bending, and rotating vibrations of the sample molecules are easier to interpret than NIR spectra and are usually modelled without filtering (Givens et al., 1997; Davis & Mauer, 2010). Not surprisingly, chemometric analysis of IR spectra may be operationally complex for newer users.

However, the combination of IR spectroscopy and chemometrics has helped to overcome most of the drawback presented by IR spectroscopy and this combination offers far quicker solutions than most other analytical techniques (Varmuza & Filzmoser, 2009). Chemometrics is employed to develop models that can relate the spectral data to the sample traits in the form of sample scores plots; where quantitative models are required, the same spectral data are modelled alongside reference data that would have been previously independently attained and ascertained by standardized laboratory methods (Peiris et al., 1999). Since the late 1990s, the development of IR spectroscopic methods has been strongly linked not only with the advance of instrument and computer technology but to a large extent to chemometrics (Omar, 2013). It has been postulated that IR spectroscopy would probably not have reached its present stage of development as fast as it did without chemometrics, as in most cases IR spectral results have been used to showcase new chemometric algorithms (Blanco & Villarroya 2002). Some could argue that chemometrics instead rides on the advances in IR spectroscopy. The fact remains that both IR spectroscopy and chemometrics have mutually benefited from

each other and in unison the techniques have provided a powerful tool to characterise a wide range of food products, including the grapevine.

2.4 Chemometrics

Chemometrics is the science of extracting chemically relevant information from chemical experiments using mathematical and statistical models (Wold, 1995). Practitioners and researchers in all applied disciplines generate several observation variables as they conduct their investigations (Varmuza & Filzmoser, 2009). This inadvertently results in the co-generation of both desirable and unwanted variables of a system; the unwanted variables have to be excluded in the final data processing. In any existing system (like in a grape berry), all the variables are naturally intertwined such that when analysed individually, they afford insignificant information about the system but when data is extracted and chemometrics is applied, the variables can be examined simultaneously to expose some underlying features of the system (Wold, 1995; Varmuza & Filzmoser, 2009). This makes chemometrics a very useful application in many areas of research including grape berry technology.

Application of chemometric methods to chemical problems was pioneered by the American chemist Thomas Isenhour in the late 1960s and many other investigators followed in his footsteps culminating in 1972 when the Swedish chemist Svante Wold coined the term *chemometrics* (Esbensen & Geladi, 1990). Since that time chemometrics has become an integral part of studying complex issues that interface chemistry/biology, chemistry/technology and chemistry/the real world (Wold, 1995; Wold & Sjöström, 1998). Chemometrics enables the joint exploration of variable contribution to a system and also determines the effect of each variable in the presence of others thus making it possible to investigate chemical variables to extract useful qualitative and quantitative essences of a given system (Wold, 1995).

Initially chemometricians were mainly interested in pattern recognition, classification, linear and non-linear mapping of chemical data but with time the thrust shifted towards understanding the properties of chemical data, and

modifying the chemometric methods to get simpler, better and more robust models (Kowalski & Bender, 1972; Wold & Sjöström, 1998). With time chemometrics grew to cover a significant range of sample features such as exploratory data analysis, supervised pattern recognition, and multivariate and or statistical analysis (Wold, 1995; Siebert, 2001).

For many years the analysis of food quality was based on univariate measurements of single parameters at a time (Martens & Naes, 1989). By the 1990s the food industry was increasingly applying chemometric techniques for research purposes (Bauer et al. 2008). The development of new multivariate techniques such as chemometrics has greatly helped to facilitate the estimation of many food quality factors through simultaneous measurement of large amounts of compositional data (Cozzolino et al., 2008).

Chemometrics has many facets that allow it to be used for both qualitative and quantitative work. Qualitative applications include principal component analysis (PCA) and discriminant analysis (orthogonal partial least squares discriminant analysis, OPLS-DA) whist qualitative applications include principal component regression (PCR), multiple linear regression (MLR) and partial least squares regression (PLS) are used for quantitative applications. The chemometrics toolbox has far many more tools than these.

PCA is at the fore of chemometrics primarily used as an exploratory tool to unmask any underlying patterns in the data variables (**X**) which emanate from the variation between the principal components in the system (Wold, 1987; Wold & Sjöström, 1998). A PCA scores plot provides a representation of sample patterns contained in a given data set and from the scores plot; the corresponding loadings plot provides information on those variables contributing to sample distribution on the scores plot (Wold & Sjöström, 1998). If no clear patterns are perceivable on the PCA scores plot, additional sample variation patterns can be explored through discriminant analysis.

Discriminant analysis is a classification technique that is used in cases where two or more groups of samples need their membership to be further constituted (Wold & Sjöström, 1998). OPLS-DA is particularly applicable for all types of classification where non-correlated systematic in **X** data needs to be analysed

(Trygg et. al, 2007). Both PCA and discriminant analysis are qualitative approaches that can give key information about a sample such as its origin, vintage, development stage, cultivar or purity. A diagnostic extension of OPLS-DA, the S-line plot is a versatile qualitative application that is better placed to show the variables that are responsible for any observed sample clustering (Wilkund et al., 2008).

Quantitative methods like PLS regression are used to deduce the concentration levels of targeted sample constituents through a defined mathematical relationship between the variable data (**X**) and independently generated reference data (**Y**) (Wold & Sjöström, 1998). Variable data may take many forms including IR spectra, NMR spectra, or chromatography measurements. The advent of chemometrics has allowed multivariate based measurements of chemical and physical properties of many products to be made (Martens & Naes, 1989). Classical analytical methods like chromatography or colorimetry use linear models based on a univariate approach that takes some optimum absorbance value to predict **Y** (Wold & Sjöström, 1998). Yet, in reality there is no single frequency where only the analyte absorbs thus chemometrics overcomes this univariate drawback through a multivariate approach that uses several or all the values of **X** to predict any given **Y** to develop a calibration model (Wold & Sjöström, 1998). In NIR and MIR spectroscopy multiple **X** values are generated at the same time making IR spectroscopy quite ideal for chemometric applications. There are several chemometric tools available for quantitative applications but PLS and is possibly the most widely used one in evaluating fruit constituents (Gómez et al., 2006).

For a calibration procedure to be considered efficient, some pertinent model properties have to be evaluated. These include the root mean square error of cross validation (RMSECV) that evaluates the model fit to the reference data; the root mean square error of estimation (RMSEE) that tests the predictive accuracy of the calibration models; the coefficient of determination (R^2) that measures the closeness of the regression line to unity (Camps & Christen, 2009). A good calibration model has the lower RMSECV and RMSEE, and, the higher R^2 (Camps & Christen, 2009).

Other statistical diagnostics are also used to test model performance and these include the root mean square error of prediction (RMSEP) that tests the model capacity to predict new samples; the ratio of prediction to deviation (RPD) that

measures the predictive performance of the model and the bias that measures the systematic differences between the predicted and reference values (Martens & Naes, 1989; Versari et al., 2008). A good model should have the lower RMSEP, low bias and high ratio of prediction to deviation (RPD > 3) and higher correlation coefficient (r^2), (Martens & Naes, 1989; Versari et al., 2008; Fernández-Novales et al, 2009). A model whose prediction accuracy is consistent and relatively immune to unknown changes in external factors is regarded as robust and desirable (Nicolaï et al., 2007).

2.5 Some applications of IR spectroscopy and chemometrics in grapevine research

The vast potential opened up by the joint application of IR spectroscopy and chemometrics has stimulated the development of several research applications that use IR spectral data to evaluate fruit quality such as grape berries. A number of research laboratories have made significant efforts to develop IR spectroscopy in tandem with chemometric investigations on the complex matrix that obtains in the grape berry.

The IR spectra in the NIR and MIR regions are composed of many well characterised functional groups that also are present in samples like the grape berry. A selected summary of some of the major structures and functional groups associated with fruit samples as detected by IR spectroscopy is given on Table 2.1.

Both qualitative and quantitative applications of IR spectroscopy on grape berry technology have been reported in the literature with the bulk of the research work seemingly leaning towards quantification of selected berry metabolites. This observation is possibly due to the fact that in most cases the targeted metabolites that attract more research interest have a major bearing on grape ripening which is the ultimate target of fruit production.

Table 2.1: A summary of the more prominent wavenumbers encountered in the NIR/MIR regions (Williams & Norris, 2001; Shenk et al, 2007; Shiroma & Rodriguez-Saona, 2009)

Wavenumber range (cm^{-1})	Functional group	Structure/Functionality
NIR region		
10220-10215	O-H first overtone	Water
8750-8745	C-H second overtone	Aromatic compounds
8600-8150	C-O stretch fourth overtone	C=O organic acids
8370-8365	C-H second overtone	R-CH_3 organic acids
8233-8227	C-H second overtone	R-C=CH_2
8165-8160	C-H second overtone	R−C≡C-H
7095-7090	O-H first overtone	R-OH
6897-6894	O-H first overtone	Water, Carbohydrates
5648-5642	C-H first overtone	-CH_2-
5617-5613	C-H first overtone	Carbohydrates
5590-5585	O-H combination	Water
MIR region		
3650-3100	O-H stretch	Water
2950-2860	C-O stretch	C=O organic acids
1612-1605	O-H stretch	Water
1497-946	C-O, C-C, C-O-H, C-O-C stretch	Fingerprint
1480-1390	C–H stretch	Organic acids
1420-1320	O-H stretch	Organic acids
1150-1060	C-O stretch	Organic acids
1065-1058	C-O stretch	Fructose
1036-1030	C-O stretch	Glucose

2.5.1. Qualitative analysis of grape berries

Qualitative analysis of berry samples using IR spectral data has been used to provide geographical, vintage, developmental stage and spoilage information of grape berries. Spectral data was also used to provide quantitative information on targeted berry metabolites.

Picque and co-workers reported successful classification of wine grapes from two geographical locations at véraison and ripe stages as part of their investigation of grape berry maturity. The MIR spectra were acquired in transmittance mode and the reference data were collected by refractometry (sugars), pH (total acidity) and titration (titratable acidity). PCA results showed that the scores plots generated from spectral data were very similar to those generated from the reference data. This result showed that spectral data can be used to provide provide quick and reliable information about grape berries in the absence of reference data. Using PCA, they were able to discriminate grapes at véraison and ripe stages from spectral data alone. Discriminant analysis (PLS-DA) was used to explore if the sources of the grape could be determined from MIR data and the results yielded a success rate of 84% in defining place of origin (Picque et al., 2010).

Barnaba and co-workers were able to separate grape berries according to vintage when they investigated the ripening of grapes over three seasons. NIR spectra were acquired in transmittance mode using a portable NIR spectrometer. Using PCA, the grapes were separated according to vintage (Barnaba et al., 2013).

Dong and co-workers used MIR spectroscopy to investigate grape spoilage in their quest to establish models that could help in the classification of grape berries according to different deterioration categories during fruit storage. MIR spectra were measured over eight days in absorbance mode from the gaseous headspace of untreated grapes in a ventilated plastic box. The results showed visual spectral differences between the berries as time and spoilage progressed. PCA scores plots categorically showed that grapes at different stages of spoilage clustered differently (Dong et al., 2014).

2.5.2 Measurement of acidity and organic acids

Organic acids are the biggest contributors to the final berry acidity and can be determined by measuring the organic acids, the titratable acidity or pH (Soyer et al., 2003; Pavloušek and Kumšta, 2011). These compounds are some of the first metabolites to be produced by the grape berry right from fruit set and, initially they are the most abundant soluble metabolites in the berry but by maturation they are the second most abundant soluble metabolites in the fruit after sugars (Coombe, 1992).

Larraín and co-workers used NIR spectroscopy to determine the acidity of selected red and white wine grapes as part of their investigation to estimate optimum grape harvest time. NIR spectra of individual *in planta* ripening berries were measured in absorbance mode with an optic field probe connected to a portable computer. The reference pH values were obtained by a digital pH-sensor. They came up with a generic model that was suitable for modelling red grape varieties and not white grapes (Larraín et al., 2008). Their results showed that the red grape varieties used in the study could be described by a single model whilst white grapes needed different models for each cultivar.

Picque and co-workers improved the scope of assessing grape berry acidity by measuring both total acidity and titratable acidity in their assessment of ripening patterns in red grapes at véraison and maturation and located at two geographical regions. The spectra were measured using transmittance MIR spectroscopy on a FT-IR Wine Scan system (Foss Corporation). Acidity reference values were determined by titration (titratable acidity) and total acidity measured by a pH probe. The highest RPD value for total acidity was 8.0 whilst for titratable acidity it was 9.2, and the models were considered good for the prediction of both total acidity and titratable acidity. Interestingly, it was observed that each geographical region required its own prediction model (Picque et al., 2010). The results showed that a global model could not be used across for different samples from different locations.

González-Caballero and co-workers further increased the scope of assessing grape berry acidity by measuring (in addition to total acidity and titratable acidity) the concentrations of individual organic acids (malic and tartaric) in white and red

grapes. Their investigation was aimed at developing a non-destructive technique for the quality control of wine grapes during on-vine ripening and on delivery to the wineries. NIR spectra were collected on individual berry samples and on the bunch in reflectance mode with a portable NIR spectrometer. Reference data for titratable acidity was determined by titration, total acidity was determined by a pH probe, tartaric acid by spectrophotometry and malic acid by reflectometry. Results derived from models for testing fruit acidity parameters yielded a screening tool sufficient to distinguish between low and high acidity values on intact grapes. A universal model was developed for predicting acidity levels for both red and white varieties (González-Caballero et al., 2010).

The results obtained in these investigations prompted the determination of a wider range of acidity markers in ripening grapes. Barnaba and co-workers used Sangiovese grapes to determine gluconic acid in addition to total acidity, titratable acidity, tartaric acid, malic acid measurements. The NIR spectra were taken in three growing seasons using a hand-held NIR spectrophotometer with an optical filter. The reference values were obtained using a FTIR Wine Scan FT 120 system. Calibration model statistics showed high R^2 values for the three acids (tartaric and gluconic acid > 0.93, malic acid > 0.86) and the RPD values for tartaric, malic and gluconic acids were 2.63, 1.19 and 1.21 respectively. This was a noteworthy outcome as it showed that even minor organic acids like gluconic acid could be determined to the same degree of applicability as major acids like malic acid. This would suggest that even the so-called minor acids can provide calibration models that can be reliably employed to measure ripening markers like acidity in grapes (Barnaba et al., 2013).

2.5.3 Measurement of sugars

Sugars are the most abundant soluble biomolecules in ripening grape berries and they provide a key indicator of grape berry quality and maturity (Coombe, 1992; Soyer et al, 2003). The content of reducing sugars such as glucose and fructose is one of the most essential parameters in grapes as it determines the eventual

alcoholic level of the wines and thus it is critical in the subsequent fermentation processes during vinification (Fernández-Novales et al, 2009).

Jarén and co-workers determined the level of soluble sugars in maturing *Viura* and *Garnacha* grapes around harvest time to help estimate harvest date in white and red grape varieties. A NIR spectrophotometer equipped with an integrating sphere in reflectance mode was used to acquire the NIR spectra. The reference values were derived from Brix sugar by refractometry. The calibration statistics were similar for the two sample types but the white variety had a higher R^2 (0.93) and lower RMSEE (1.04 ^{o}Brix) than the red variety (R^2 = 0.89 and RMSEE =1.05 ^{o}Brix). They concluded that regarding the estimation of ripening, each grape variety needed to be modelled separately (Jarén, et al, 2001).

Herrera and co-workers used grape soluble sugars as a ripening index for a variety of red and white grapes in an attempt to estimate grape berry ripening times. A portable NIR spectrometer in both transmittance and interactance modes was used for the NIR spectral acquisitions and the reference Brix values were obtained by refractometry. The calibration coefficients showed that the transmission mode produced better models for some varieties, with interactance mode proving to be better for others (Herrera et al., 2003). These results showed that in some cases more than one spectral acquisition technique may be needed to get more informative results.

Fernández-Novales and co-workers determined reducing sugar content in an investigation that was aimed at controlling grape quality parameters for red and white berry varieties during ripening and at harvest. NIR spectra were measured in transmittance mode with a fiber spectrometer and sugar reference values determined by titration. Four common fingerprint spectra were found for the determination of reducing sugars content in the grapes irrespective of the variety (Fernández-Novales et al., 2009). This was an interesting outcome as in most cases the best NIR models are obtained by using the full spectral range.

Expounding on their assessment of sugars González-Caballero and co-workers collected spectra on grape samples at bunch and at individual berries making up the bunch as part of their assessment of changes in internal quality properties of wine grapes during on-vine ripening and at harvest. NIR spectra were acquired in

reflectance mode using an NIR spectrometer. The reference sugar values were determined by refractometry. Results derived from the PLS models for testing sugar concentrations yielded very high R^2 (> 0.93) for both bunch and individual berry samples (González-Caballero et al., 2010). The results gave credence to the notion that it is possible to offer faster assessment of grape batches at delivery points by assessing bunches only.

Whilst investigating the ripening patterns in red grapes at véraison and maturation stages from two geographical regions, Picque and co-workers measured sugars with a FOSS Wine Scan system in MIR transmittance spectroscopy. The group analyzed a wide range of oenological parameters including sugars whose reference data were determined by refractometry. It was found that of all the parameters that were investigated, sugars afforded the best PLS models: average R^2 > 0.88 and RPD > 3.33 compared to any other metabolite. More importantly, the results showed that a universal model for sugar determination can be used for both locations (Picque et al., 2010).

Barnaba and co-workers were one of the first groups to measure the individual levels of glucose and fructose in addition to Brix using NIR spectroscopy in their investigation of the ripening of *Sangiovese* grapes. NIR spectra were acquired by a portable NIR-AOTF spectrophotomer in diffuse reflectance mode. The reference values were determined by refractometry. In comparison to other oenological factors, fructose, glucose and total sugars had the highest R^2 values (0.92, 0.93 and 0.94) and RPD values (2.51, 2.69 and 3.03) (Barnaba et al., 2013). Even though they worked with ripening berries where the levels of sugars are always high and more accurate to determine, their results showed that evaluating sugars could be one of the most reliable ways of investigating biomarkers of grape berry ripening.

2.5.4 Measurement of phenolics, anthocyanins and tannins

Phenolic compounds are plant secondary metabolites with phenolics in grape showing a vast structural diversity ranging from simple molecules to polyphenols (Lorrain et al., 2013). Secondary metabolism products such as polyphenols are

increasingly becoming more important for advanced definition of berry quality and the concept of phenolic ripening is now a significant maturation indicator of grape berries as the content of phenolic compounds underpins the state of phenolic ripening (Fragoso et al., 2006).

Fragoso and co-workers used MIR spectroscopy to devise a fast measurement protocol for phenolic ripening. Six fresh red grape varieties at different phenolic ripening stages were sampled to accommodate the high natural variability of grapes when building the calibration models. The MIR spectra were acquired in reflectance mode and the total phenolic compounds, anthocyanins and tannins used as reference values were acquired through UV-visible spectroscopy. PLS regression analysis gave $R^2 > 0.91$ for the three sets of compounds, RPDs ranged from 2.6 – 7.1 and RMSEPs ranged from 4.3 – 8.0%. These results showed that MIR spectroscopy could be reliably used to monitor phenolic ripening in red grapes at harvest time (Fragoso et al., 2006).

During their investigation of the monitoring of phenolic ripening in red grapes, Ferrer-Gallego and co-workers used NIR spectroscopy to measure the phenolic acids, anthocyanin and flavanol concentrations of grape skins and intact grapes. NIR spectra were acquired in diffuse reflectance mode and the reference data were obtained by HPLC. For the whole range of phenolic substances the statistics were deemed good for prediction (R^2 = 0.77 – 0.94; RPD = 2.1 – 4.8). External validation showed that the best results for the determination of flavonols were obtained with intact grapes as compared to using grape skins Intact grapes had a deviation of 7.8% (taken as the deviation between reference and spectral data) whilst skins had a deviation of 10.7%. Good results in the external validation process were also obtained for the determination of total phenolic compounds with a deviation of 11.7% using intact grapes and 14.7% using grape skins. It was evident that better results were obtained from the direct recording of intact grape as compared to skin only (Ferrer-Gallego et al., 2011). The results showed the importance of sample selection when carrying out such advanced investigation as phenolic ripening.

Picque and co-workers measured total phenols and anthocyanins in their assessment of ripening patterns in red grapes at véraison and maturation at two

geographical regions. They acquired the spectra by transmittance MIR spectroscopy on a FOSS Winescan system. Reference values for phenolics and anthocyanins were determined by filtration/centrifugation. PLS regression analysis showed R^2 values of 0.51 and less (average 0.39), highest RPD value of 1.5 (average 1.13) for phenolics. For anthocyanins the highest R^2 was 0.73 (average 0.44), highest RPD value of 2.1 (average 1.33). The anthocyanin models were deemed to be not satisfactory for predicting berry ripening (Picque et al., 2010).

Barnaba and co-workers determined phenolics and anthocyanins in their assessment of grape ripening. The NIR spectra were acquired in transmittance mode, reference data generated through FTIR Wine Scan FT 120 facility. PLS regression analysis showed R^2 values of 0.69 and 0.76 for phenolics and anthocyanins respectively; RPDs were less than 1.35 for both sets of compounds. The phenolic and anthocyanin models were deemed unsuitable for ripening prediction (Barnaba et al., 2013).

Some differences in model statistical details between the two NIR techniques by the Barnaba et al (2013) and Ferrer-Gallego et al. (2011) were interesting. Whilst the R^2 values were fairly similar the RPD values were much higher in the latter. The quality of reference data was reported to be a major source of error that can result in the distortion of a PLS model's performance if the data are not accurate enough (Smyth & Cozzolino, 2011). Ferrer-Gallego et al., (2011) used HPLC reference data whereas Barnaba et al., (2013) used reference data from supernatants that were analysed by an FTIR Wine Scan FT 120. This discrepancy in regard to reference data was also observed when the anthocyanins and phenolics were determined by MIR spectroscopy. Fragoso et al. (2011) used reference data acquired through UV-Visible spectroscopy whilst Pique et al. (2010) used filtration/centrifugation techniques to get reference data. These variations appear to support the observation by Smyth & Cozzolino (2011) that in order to construct good PLS calibration models the reference data are better acquired by analytical grade standards. Considering the amount of work that is put into model development, it is worth it to use very reliable reference data.

2.5.5 Measurement of chemical markers to determine grape berry infection

An exciting application of IR spectroscopy and chemometrics was show-cased in a pioneering project seeking to monitor grape pathogens like *Botrytis cinerea*. This fungal infection is traditionally detected through microscopy, immunological techniques, visual inspection and molecular biology as routine quality control measures to assess the sanitary status of grapes before crushing.

Versari and co-workers quantified gluconic acid and glycerol as chemical markers to determine *B. cinerea* infection on harvested grapes (Versari et al., 2008). MIR spectra were acquired using a FOSS GrapeScan instrument in MIR transmittance spectroscopy with the reference data generated by HPLC. PLS regression analysis showed the best results were obtained for gluconic acid (R^2 = 0.98, RMSECV = 0.63 and RPD = 7.0). This FT-MIR technique was deemed as suitable for process control and good enough for routine analysis to inspect the botrytisation of grapes (Versari et al., 2008).

2.5.6 Measurement of grapevine water potential

De Bei and co-workers used NIR spectroscopy to measure grapevine water potential in three grapevine varieties. NIR spectra were acquired using a mobile spectrometer in diffuse reflectance mode on both sides of selected leaves and the reference values were obtained by a pressure chamber on the same leaves where spectra were taken. Good calibration models were obtained: R^2 = 0.67-0.84, RMSECV =0.18-0.26 MPa for Chardonnay vines; R^2 = 0.83-0.92, RMSECV = 0.08-0.09 MPa for Cabernet Sauvignon vines and R^2 = 0.84-0.92, RMSECV = 0.07-0.11 MPa for Shiraz vines. Additional data obtained from glasshouse and field studies of Cabernet Sauvignon and Shiraz vines were used to develop a global calibration model for the prediction of water potential. A high R^2 (0.87) and a low RMSECV (0.1 MPa) were observed. The resultant global calibration for the two varieties was taken as a good indicator towards the future estimation of water potential as it was postulated that a universal calibration that would be able to predict water potential

for all grapevine varieties in different environments could be constructed (De Bei et al., 2011).

2.6 Summary of reviewed articles

In this review it was clear that both NIR and MIR spectroscopy have been used to measure many oenological parameters of grape berry ripening. Chemometric tools ranging from PCA, discriminant analysis (PLS-DA) and PLS were used by different researchers in their investigations.

From the various sets of results it was observed that primary grape berry metabolites especially sugars and organic acids produced models that were generally considered to be reliable for predicting grape berry maturation. Secondary metabolites like phenolics and anthocyanins gave a mixed bag of outcomes: with some groups coming up with models that were considered good enough for prediction of maturation and other groups coming up with models that were considered unsuitable for predictive purposes.

It was also noted different research groups did not always come to the same conclusions in regard to the prediction models even in cases were the same oenological parameters were investigated in the same grape varieties. Red grape varieties in most cases were reported to be described by a universal model in cases where the samples were sourced from within the same region but when different locations were used, different models were required for predictive purposes. With white grape varieties it was generally observed that each cultivar would need its own model. Based on these results, the idea of a global grape model that would be used for all cases seems to be unattainable from the IR spectroscopic approach used and more work would need to be carried out to look into the variations observed.

Product authenticity and quality are key issues in the grape and wine industries as modern consumers are increasingly becoming more conscious of the occurrence of counterfeits that are competing with such important products as fruit and wine (Bevin et al, 2006). The summary of results showed that significant qualitative information like source of grapes, developmental stage, vintage or freshness could

be obtained from IR spectral data alone. Such information can be used by producers to keep reliable sample databases such as records of fruit development and product source to validate product authenticity especially in the wine sector where vintage plays such a crucial part in purchasing decisions.

The study on measurement of water potential in grapevines demonstrated that IR spectroscopy can provide a fast and reasonable assessment of leaf water potential in the field and this study can be used for example in scheduling irrigation.

2.7 Cited literature

Adato, R., & Altug, H. (2013). In-situ ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas. *Nature Communications,* 1-14.

Alexandersson, E., Jacobson, D., Vivier, M.A., Weckwerth, W., & Andreasson, E. (2014). Field-omics – understanding large-scale molecular data from field crops. *Frontiers in Plant Science, 5,* 1-6.

Ali, K., Maltese, F., Fortes, A.M., Pais, M.S., Choi, Y.H., & Verpoorte, R. (2011). Monitoring biochemical changes during grape berry development in Portuguese cultivars by NMR spectroscopy. *Food Chemistry. 124,* 1760–1769.

Arana, I., Jarén, C., & Arazuri, S. (2005). Maturity, variety and origin determination in white grapes (Vitis vinifera L.) using near infrared reflectance technology*. Journal of. Near Infrared Spectroscopy, 13,* 349-357.

Arzamastsev, A.P., Sadchikova, N.P., & Titova, A.V. (2008). Structure of chemical compounds, methods of analysis and process control. *Pharmaceutical Chemistry Journal, 42,* 466-470.

Barnaba, F.E., Bellincontro, A., & Mencarelli, F. (2013). Portable NIR-AOTF spectroscopy combined with winery FTIR spectroscopy for an easy, rapid, in-field monitoring of Sangiovese grape quality. *Journal of Agricultural and Food Chemistry, 94,* 1071-1077.

Barth, A., & Haris, P. (2009). Infrared spectroscopy – past and present. In Barth, A. & Haris, P. (Eds), *Biological and Biomedical Infrared Spectroscopy* (pp. 1-52). Amsterdam: IOS Press.

Bevin, C.J.; Fergusson, A.J.; Perry, W.B.; Janik, L.J., & Cozzolino D.(2006). Development of a rapid "fingerprinting" system for wine authenticity by mid-infrared spectroscopy. *Journal of Agricultural and Food Chemistry, 54,* 9713-9718.

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: a rapid-response analytical tool*. Trends in Analytical Chemistry, 21,* 240-250.

Brügel, W. (1962). *Infrared Spectroscopy*. London: Methuen & Co Ltd*,* (Chapter 7 Part I, Chapter 3 Part III).

Butz, P., Hofmann, C., & Tauscher, B. (2005). Recent developments in noninvasive techniques for fresh fruit and vegetable internal quality analysis. *Journal of Food Science. 70,* 131-141.

Camps, C., & Christen, D. (2009). Non-destructive assessment of apricot fruit quality by portable visible-near infrared spectroscopy. *Food Science and Technology, 42,* 1125–1131.

Conde, C., Silva, P., Fontes, N., Dias, A.C.P., Tavares, R.M., Sousa, M.J., Agasse, A., Delrot, S., & Gerós, H. (2007). Biochemical changes throughout grape berry development and fruit and wine quality. *Food,.* 1-18.

Coombe, B.G. (1992). Honorary research lecture: Research on the development and ripening of the grape berry. *American Journal of Enology and Viticulture 43,* 101-110.

Coombe, B.G. (1995). Adoption of a system for identifying grapevine growth stages. Australian *Journal of Grape and Wine Research, 1,* 100-110.

Coombe, B.G., & McCarthy, M.G. (2000). Dynamics of grape berry growth and physiology of ripening*. Australian Journal of Grape and Wine Research 6,* 131–135.

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,* 7631-7636.

Dal Santo, S., Tornielli, G.B., Zenoni, S., Fasoli, F., Farina, L., Anesi, A. Guzzo, F., Massimo Delledonne, M., & Pezzotti, M. (2013). The plasticity of the grapevine berry transcriptome. *Genome Biology, 14,* 1-18.

Davis, R., & Mauer, L.J. (2010). Fourier transform infrared (FT-IR) spectroscopy: A rapid tool for detection and analysis of foodborne pathogenic bacteria. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology,* 1582-1594.

De Bei, R., Cozzolino, D., Sullivan, W., Cynkar, W., Fuentes, S., Dambergs, R., Pech, J, & Tyerman, S. (2011). Non-destructive measurement of grapevine water potential using near infrared spectroscopy. *Australian Journal of Grape and Wine Research 17,* 62-71.

Dong, D., Zheng, W., Wang, W., Zhao, X., Jiao, L., & Zhao, C. (2014). Analysis and discrimination of grape spoilage via volatiles: a comparison between long optical path Fourier-transform-infrared spectroscopy and sensor arrays. *The Analyst, 139,* 5028-5034.

Dufour, E. (2009). Principles of infrared spectroscopy. In: Sun, D-W. (Ed), *Infrared spectroscopy for food quality analysis and control* (pp.3-27). Boca Raton: CRC Press.

Esbensen, K., & Geladi, P. (1990). The start and early history of chemometrics: Selected interviews, Part 2. *Journal of Chemometrics, 4,* 389-412.

Fernández-Novales, J., López, M-I., Sánchez, M-T., Morales, J., & González-Caballero, V. (2009). Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines. *Food Research International, 42,* 285-291.

Ferrer-Gallego, R.; Hernández-Hierro, J.M.; Rivas-Gonzalo, J.C., & Escribano-Bailón, T.M. (2011). Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *LWT - Food Science and Technology, 44,* 847-853.

Foley, W.J., McIlwee, A., Lawler, I., Aragones, L., Woolnough, A.P., & Berding, N. (1998). Ecological applications of near infrared reflectance spectroscopy – a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. *Oecologia, 116,* 293-305.

Fragoso, S., Aceña, L., Guasch, J. Busto, O., & Mestres, M. (2011). Application of FT-MIR spectroscopy for fast control of red grape phenolic ripening. *Journal of Agricultural and Food Chemistry, 59,* 2175-2183.

García de Cortázar-Atauri, I., Brisson, N., & Gaudillere, J.P. (2009). Performance of several models for predicting budburst date of grapevine (Vitis vinifera L*.). International Journal of Biometeorology, 53,* 317-326.

Gholami, M., Hayasaka, Y., Coombe, B.G., Jackson, J.F., Robinson, S.P., & Williams, P.J., (1995). Biosynthesis of flavour compounds in Muscat Gordo Blanco grape berries. *Australian Journal of Grape and Wine Research, 1,* 19-24.

Gishen, M.; Dambergs, R.G., & Cozzolino, D. (2005). Grape and wine analysis – enhancing the power of spectroscopy with chemometrics:A review of some applications in the Australian wine industry. *Australian Journal of Grape and Wine Research 11,* 296-305.

Givens, D. I., De Boever, J. L., & Deaville, E. R. (1997). The principles, practices and some future applications of near infrared spectroscopy for predicting the nutritive value of foods for animals and humans. *Nutrition Research Reviews, 10,* 83-114.

Gómez, A.H., He, Y., & Pereira, A.G. (2006). Non-destructive measurement of acidity, soluble solids and firmness of Satsuma mandarin using Vis/NIR-spectroscopy techniques. *Journal of Food Engineering, 77,* 313–319.

González-Caballero, V., Sánchez, M-T., López, M-I. & Pérez-Marín, D. (2010). First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery. *Journal of Food Engineering, 101,* 158-165.

Herrera, J., Guesalaga, A. & Agosín, E. (2003). Shortwave near infrared spectroscopy for non-destructive determination of maturity of wine grapes. *Measurement Science and Technology, 14,* 689-697.

Huang, H., Yu, H., Xu, H., & Ying, Y. (2008). Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. *Journal of Food Engineering, 87*, 303–313.

Hunter, J.J., Visser, J.H., & De Villiers, O.T. (1991). Preparation of grapes and extraction of sugars and organic acids for determination by high performance liquid chromatography. *American Journal of Enology and Viticulture, 42,* 237-244.

Jarén, C., Ortuño, C., Arazuri, S., Arana, J.I., & Salvadores, M.C. (2001). Sugar determination in grapes using NIR technology. *International Journal of Infrared and Millimeter Waves, 22,* 1521-1530.

Kelly, J.F.D., Downey, G., & Fourtier, V. (2004). Initial study of honey adulteration by sugar solutions using mid-infrared (MIR) spectroscopy and chemometrics. *Journal of Agricultural and Food Chemistry, 52,* 33-39.

Kowalski, B.R., & Bender, C.F. (1972). Pattern recognition. A powerful approach to interpreting chemical data. *Journal of the American Chemical Society, 94,* 5632-5639.

Larraín, M., Guesalaga, A.R., & Agosín, E. (2008). A multipurpose portable instrument for determining ripeness in wine grapes using NIR spectroscopy. *IEEE Transactions on Instrumentation and Measurement, 57,* 294-502.

Lorrain, B, Ky, B., Pechamat,L., & Teissedre, P-L. (2013). Evolution of analysis of polyhenols from grapes, wines,and extracts. *Molecules, 18*, 1076-1100.

Muñoz-Robredo, P., Robledo, P., Manríquez, D., Rosa Molina, R., & Defilippi, B.G. (2011). Characterization of sugars and organic acids in commercial varieties of table grapes. *Chilean Journal of Agricultural Research, 71,* 452-458.

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I. & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46,* 99-118.

OIV Statictical Report on World Vitiviniculture (2013). *Grape,* 11-13.

Omar, A.F. (2013). Spectroscopic profiling of soluble solids content and acidity of intact grape, lime, and star fruit. *Sensor Review, 33,* 238–245.

Pavloušek, P., & Kumšta, M. (2011). Profiling of primary metabolites in grapes of interspecific grapevine varieties: sugars and organic acids. *Czech Journal of Food Sciences 29,* 361-372.

Peiris, K. H. S.; Dull, G. G., Leffler, R. G., & Kays, S. J. (1999). Spatial variability of soluble solids or dry-matter content within individual fruits, bulbs, or tubers: Implications for the development and use of NIR spectrometric techniques. *HortScience, 34,* 114-118.

Picque, D., Lieben, P., Chrétien, Ph., Beguin, J., & Guerrin, L. (2010). Assessment of maturity of Loire Valley wine grapes by mid-infrared spectroscopy. *Journal International des Sciences de la Vigne et du Vin 44,* 219-229.

Quilitzsch, R., Baranska, M., Schulz, H., & E. Hoberg, E. (2005). Fast determination of carrot quality by spectroscopy methods in the UV-VIS, NIR and IR range. *Journal of Applied Botany and Food Quality 79,* 163-167.

Reynolds, A.G. (2010). Viticultural and vineyard management practices and their effects on grape and wine quality. In Reynolds, A.G.(Ed), *Managing Wine Quality, Vol. 1: Viticulture and Wine Quality* (pp. 365-429). Boca Raton: CRC Press.

Rolle, L., Torchio, F., Giacosa, S., Río Segade, S., Cagnasso, E. & Gerbi, V. (2012). Assessment of physicochemical differences in Nebbiolo grape berries from different production areas and sorted by flotation. *American Journal of Enology and Viticulture 63,* 195-205.

Shenk, J.S., Workman Jr., J.J., & Westerhaus, M.O. (2008). Application of NIR spectroscopy to agricultural products. In: Burns, D.A., & Ciurczik E.W. (Eds), *Handbook of NIR Analysis* (pp. 349-360). New York: CRC Press.

Shiroma, C., & Rodriguez-Saona, L. (2009). Application of NIR and MIR spectroscopy in quality control of potato chips. *Journal of Food Composition and Analysis, 22,* 596–605.

Simon, V. (1966). *Infrared Radiation.* Princeton: D. Van Nostrand Company, Inc. (Chapter 1).

Sinelli, N., Spinardi, A., Di Egidio, V., Mignani, I., & Casiraghi, E. 2008. Evaluation of quality and nutraceutical content of blueberries (*Vaccinium corymbosum L.*) by near and mid-infrared spectroscopy. *Postharvest Biology and Technology, 50,* 31-36.

Smyth, H.E., & Cozzolino, D. (2011). Applications of infrared spectroscopy for quantitative analysis of volatile and secondary metabolites in plant materials. *Current Bioactive Compounds, 7,* 66-74.

Soyer, Y., Koca, N. & Karadeniz, F. (2003). Organic acid profile of Turkish white grapes and grape juices. *Journal of Food Composition and Analysis 16,* 629-636.

Teixeira, A.H.C., Tonietto, J., Pereira, G.E., & Hernandez, F.B.T. (2014). Characterization of the wine grape thermohydrological conditions in the tropical Brazilian growing region: long-term and future assessments. *ISRN Agronomy, 14,* 1-14.

Topalovic, A., & Mikulic-Petkovsek, M. (2010). Changes in sugars, organic acids and phenolics of grape berries of cultivar Cardinal during ripening. *Journal of Food, Agriculture and Environment, 8,* 223-227.

Trygg, J., Holmes, E. & Lundstedt, T. (2007). Chemometrics in metabonomics. *Journal of Proteome Research, 6,* 469-479.

Varmuza, K., & Filzmoser, P. (2009). Introduction; Principal Component Analysis. In: Varmuza, K., & Filzmoser, P. (Eds), *Introduction to Multivariate Statistical Analysis in Chemometrics* (pp. 1-101). New York: CRC Press).

Versari, A., Parpinello, G.P., Mattioli, A.U., & Galassi, S. (2008). Determination of grape quality at harvest using Fourier-transform mid-infrared spectroscopy and multivariate analysis. *American Journal of Enology and Viticulture, 59,* 317-322.

Walker, H.J. (2000). A brief history of infrared spectroscopy. *Infrared Spectroscopy, 41,* 510-513.

White, P. J. 2002. Recent advances in fruit development and ripening: an overview. *Journal of Experimental Botany, 53,* 1995-2000.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E.J., Edlund, U., Shockcor, J.P., Gottfries, J., Moritz, T., & Trygg, J. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry, 80,* 115-122.

Williams, P., & Norris, K. H. (2001). Variables affecting near infrared spectroscopic analysis. In Williams, P., & Norris, K.H. (Eds). *Near infrared technology in the agriculture and food industries* (pp. 171–185). St Paul: The American Association of Cereal Chemists.

Wold, S. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2,* 37-52.

Wold, S. (1995). Chemometrics; what do we mean with it, and what do we want from it? *Chemometrics and Intelligent Laboratory Systems, 30,* 109-115.

Wold, S., & Sjöström, M. (1998). Chemometrics, and its roots in physical organic chemistry. *Acta Chemica Scandinavica, 52,* 517–523.

CHAPTER 3

A rapid qualitative and quantitative evaluation of grape berries at various stages of development using Fourier-transform infrared spectroscopy and multivariate data analysis

This chapter was submitted for publication in the Food Chemistry and has multiple authors. The contributions of the co-authors are described in the preface.

A rapid qualitative and quantitative evaluation of grape berries at various stages of development using Fourier-transform infrared spectroscopy and multivariate data analysis

Davirai M. Musingarabwi, Hélène H. Nieuwoudt, Philip R. Young, Hans A. Eyéghè-Bickong, Melanè A. Vivier*

Institute for Wine Biotechnology, Department Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

*Author to whom correspondence should be addressed (e-mail mav@sun.ac.za tel +27-21-8083770 fax +27-21-8083771)

Highlights

- IR techniques are described for qualitative and quantitative profiling of grapes
- IR techniques are robust, rapid and accurate and require minimal sample processing
- Accurate prediction of individual sugars and organic acid concentrations using MIR
- Methods are applicable to grape berries throughout development

3.1 Abstract

Fourier transform (FT) near-infrared (NIR) and attenuated total reflection (ATR) FT mid-infrared (MIR) spectroscopy were used to qualitatively and quantitatively analyse *Vitis vinifera* L. cv Sauvignon blanc grape berries. FT-NIR and ATR FT-MIR spectroscopy, coupled with spectral preprocessing and multivariate data analysis (MVDA), provided reliable methods to qualitatively assess berry samples at five distinct developmental stages: green, pre-véraison, véraison, post-véraison and ripe (harvest), without any prior metabolite extraction. Compared to NIR spectra, MIR spectra provided more reliable discrimination between the berry samples from the different developmental stages. Interestingly, ATR FT-MIR spectra from fresh homogenized berry samples proved more discriminatory than spectra from frozen homogenized berry samples. Different developmental stages were discriminated by principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA). In order to generate partial least squares (PLS) models from the MIR/NIR spectral datasets; the major sugars (glucose and fructose) and organic acids (malic acid, succinic acid and tartaric acid) were separated and quantified by high performance liquid chromatography (HPLC) and the data used as a reference dataset. PLS regression was used to develop calibration models to predict the concentration of the major sugars and organic acids in the berry samples from different developmental stages. Our data show that Infrared (IR) spectroscopy could provide a rapid, reproducible and cost-effective alternative to the chromatographic analysis of the sugar and organic acid composition of grape berries at various developmental stages, using small sample volumes and requiring limited sample preparation. This provides scope and support for the possible development of hand-held devices to assess quality parameters in field-settings in real-time and non-destructively using IR technologies.

Keywords: Sauvignon blanc, grape berry development, ATR FT-MIR, FT-NIR, MVDA, PCA, OPLS-DA, PLS

3.2 Introduction

The grape berry has been described as a complex biochemical factory, that produces a diverse array of primary (*e.g.* sugars, organic acids) and secondary metabolites (*e.g.* phenolic and aromatic compounds) that can occur at various developmental and/or maturation stages (Coombe & McCarthy, 2000; Gholami et al, 1995). The metabolism of these compounds in the berries is typically used as indicators of the progression of development and ripening (Coombe, 1992). Since the concentration of the primary metabolites (i.e. sugars and organic acids) at ripeness contributes to the perceived quality of grapes and the subsequent wine (Rusjan et al., 2008), their relative abundance is typically used to direct operational decisions in the vineyard that include vineyard manipulation, harvest dates and the desired wine style(s).

The study of these metabolites in grape berries is, however, complicated by the inherent heterogeneity found in vineyards (or any field-grown crop). Grape berry development and ripening has been shown to be asynchronous, with heterogeneity found on multiple levels: (1) between berries on a single bunch; (2) between bunches on a vine; (3) between vines in a vineyard; and (4) between vineyards in a region (Jackson and Lombard, 1993; Dal Santo et al., 2013). To account for this heterogeneity, profiling and classification methods are usually employed to analyse the major grape berry metabolites (García de Cortázar-Atauri et al., 2009; Šuklje et al., 2012). It is clear that the existence of variability has a significant impact on grape berry sampling (and relevance of subsequent data generated), leading to the necessity for updated workflows and technologies to account for this variability (Alexandersson et al., 2014).

Classical analytical methods such as high performance liquid chromatography (HPLC), gas chromatography (GC) and colorimetry have been extensively used for the quantitative analysis of grape berry composition. These "traditional" methods though lengthy and expensive are very accurate and still remain the gold standards of high end research applications (Herrera et al., 2003), but do not satisfy the demands of the modern global grape and wine industries. The fresh fruit industry requires methods that allow rapid, reliable and relatively low-cost analyses and/or non-invasive techniques to estimate or profile key metabolites (biomarkers) (González-Caballero et al., 2010). Moreover, the prediction of metabolites such as sugars and organic acids can assist in the implementation of appropriate management strategies to control

desired grape attributes (*e.g.* size, colour, acidity, chemical composition) (Ali et al., 2011).

Spectroscopic techniques, such as nuclear magnetic resonance (NMR) and infrared (IR) in combination with multivariate data analysis (MVDA), have found increasing application in agriculture (Gishen et al., 2005; Quilitzsch et al., 2005). NMR spectroscopy is a non-selective, simple and relative rapid technique ideal for the profiling of a broad range of metabolites such as those found in grape berries (Ali et al, 2010). It has been successfully applied for the estimation of sugars and some organic acids in berries in a number of Portuguese grape cultivars (Ali et al., 2011). NMR spectroscopy, however, requires sample preparation and use of deuterated solvents for analysis. IR spectroscopy is suitable for the determination of compounds containing polar functional groups such as -OH, C-O, and N-H (Blanco & Villaroya, 2002) and offers an alternative to estimate sugars and organic acids (polar compounds) present in grape berries without prior analyte extraction.

The development of FT-IR spectroscopy instrumentation combined with MVDA techniques has given impetus to the application of rapid methods for predicting the presence and/or concentration of specific chemical constituents in fruit and vegetable matrices (Nicolaï et al., 2007). IR spectroscopy in the NIR (12000-4000 cm^{-1}, 833-2500 nm) and the MIR (4000-400 cm^{-1}, 2500-25000 nm) spectral regions (Kelly et al., 2004) has been extensively applied as alternative analytical tools in the food industry (Bauer et al., 2008). NIR spectroscopy is attractive due to the fact that it is rapid and non-destructive and it has the additional benefit of penetrating finite distances (centimeters) in fruit (Nicolaï et al., 2007). This versatility makes NIR spectroscopy an alternative to conventional extractive analyses to be used for the acquisition of surface and internal characteristics of foods. In comparison to NIR spectroscopy, MIR spectroscopy provides spectral information that is easier to interpret due to the higher resolution of the fundamental MIR vibrational absorption bands compared to the overtone and combination NIR absorption bands (Kelly et al., 2004). With short runtimes (seconds), IR spectroscopy is not only rapid, but has relatively low operational costs and since no chemical waste is generated, is comparatively environmentally friendly (Nicolaï et al., 2007; González-Caballero et al., 2010). Portable (hand-held) visible/NIR spectroscopic devices are currently commercially available and provide a relatively simple, non-destructive analysis and have been used to estimate, for example, berry ripeness in the Italian red grape cultivar Nebbiolo (Giovenzana et al., 2014). However, no single technique can solve

all analytical problems and IR spectroscopic techniques are no exception. Direct IR spectroscopy cannot be relied upon to measure substances occurring below the mg/g level (Bauer et al., 2008). In aqueous measurements (prevalent in most fruit products), water presents a major obstacle for IR analysis as the water O-H bending absorption can mask most other sample signals to the level that they may appear non-existent (Adato & Altug, 2013). Direct interpretation of IR spectra may be arduous especially in the NIR region, as some of the absorption peaks are emanating from complex overtone and high frequency combinations from primary absorptions (Kelly et al., 2004).

NIR spectrocopy has been successfully implemented to monitor a number of metabolites in grape berries at post-véraison/ripe developmental stages that include: sugars (Jarén et al., 2001, Giovenzana et al., 2014), organic acids (González-Caballero, 2010) and anthocyanins (Cozzolino et al., 2005). MIR spectroscopy has been used to study phenolic compounds in post-véraison/ripe berries (Fragoso et al., 2011). Fruit quality, however, is dependent on the compositional changes occurring throughout development due to the cumulative effects of the environment (amongst other factors) on the fruit (Alexandersson et al., 2014). To our knowledge, there is no reported use of FT-NIR and FT-MIR spectroscopy in grape berry analyses over the entire developmental process. An approach that incorporates/utilizes the advantages of both NIR and MIR spectroscopic techniques could potentially provide novel insights into grape berry development and/or stage-specific grape berry responses. The ability to assess berry composition using high-throughput technologies would support the practical implementation of comprehensive analysis of developing and ripening berries to provide a qualitative and/or quantitative profile for a vineyard (Rolle et al., 2012; Alexandersson et al., 2014).

In this study the potential of NIR and MIR spectroscopy to assess grape berry development qualitatively (via PCA and discriminant analysis) and quantitatively, (via PLS regression models derived from reference data sets), was investigated. The NIR and MIR spectra were collected from grape berries sampled at five distinct developmental stages: green, pre-véraison, véraison, post-véraison and ripe/harvest stages. Qualitative analysis was carried out using MVDA of the respective spectra. HPLC was used to generate reference analytical data for the major sugars (glucose and fructose) and organic acids (tartaric acid, malic acid, succinic acid) for quantitative analysis. The sugars and organic acids were selected as they are well described biomarkers of berry growth and development (Coombe, 1992; Topalovic &

Mikulic-Petkovsek, 2010). PLS regression analysis (Wold et al., 2001; Wilkund et al., 2008) of the spectra was used to develop calibration models for the prediction of the concentrations of these metabolites in the grape berry at five distinct developmental stages.

3.3 Materials and methods

High performance liquid chromatography and IR (NIR and MIR) spectroscopy were used in this study. Fig. 3.1 provides a flow chart for an overview of the respective steps and how they are related: (1) sample collection, (2) sample processing, and (3) sample analysis, (4) data analysis, and (5) outputs derived from the data. Experimental details are provided in the relevant sections of the materials and methods.

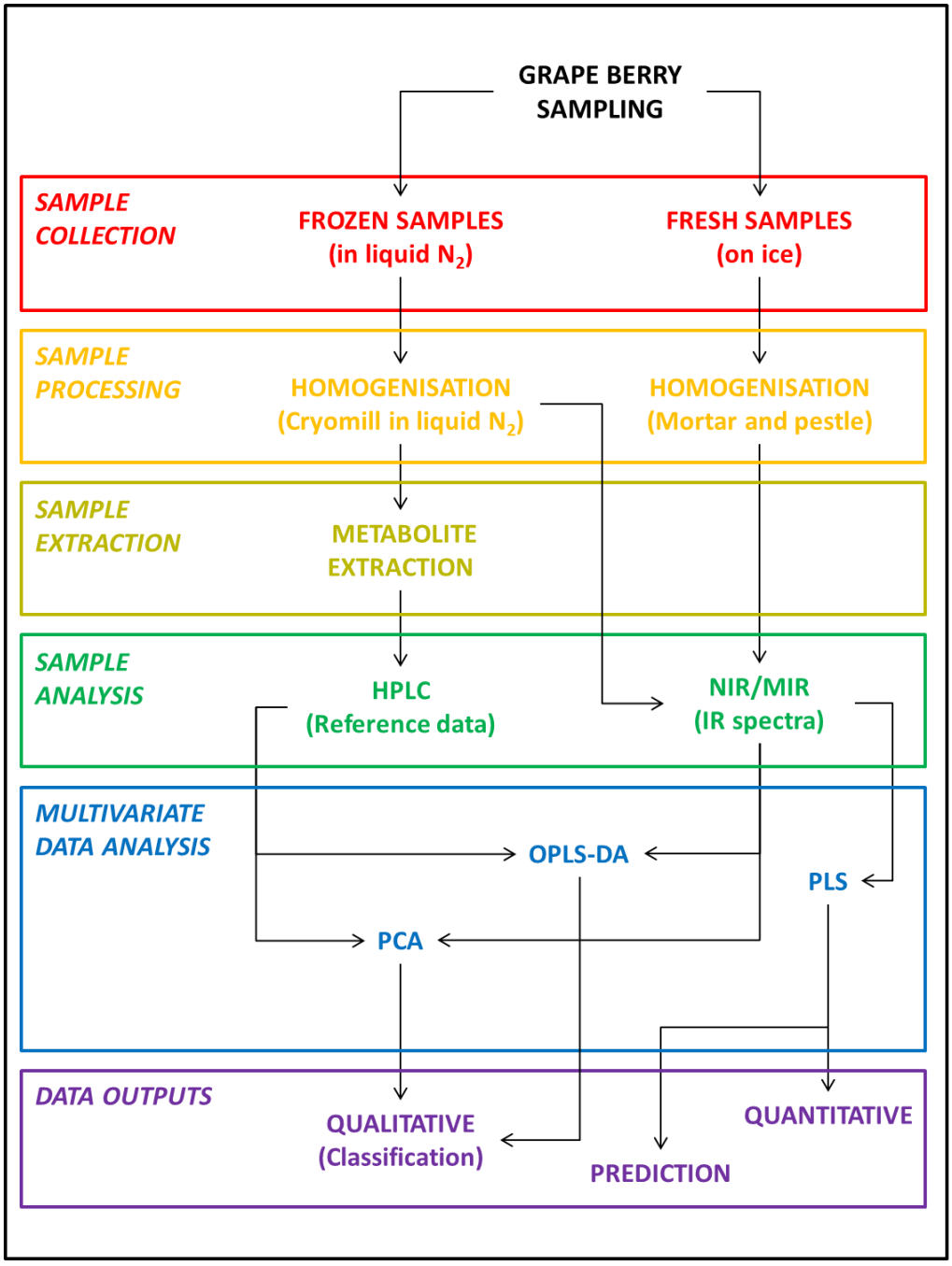


Figure 3.1: Flowchart of methodologies applied for grape berry analysis using chromatography (as reference data) and IR spectroscopy

3.3.1 Berry sampling and processing

Grape berries from *Vitis vinifera* L. cv. *Sauvignon blanc* (clone 316 grafted to 101-14 Mgt), were established in the Elgin wine region (Overberg, coastal region of the Western Cape, South Africa: 34° 9'52.19"S; 19° 0'57.48"E) in 2004. The vines have a NW-SE row orientation with 2.5 m × 1.8 m spacing. Shoots were positioned vertically (vertical shoot positioning trellis system) with two levels of movable wires and pruned using a double cordon system with two buds per spur. Irrigation was applied to avoid

water constraints to the experimental vines. Two rows consisting of six panels (four vines per panel) were sampled at five developmental stages: green (Eichhorn-Lorenz (E-L) system stage 32), pre-véraison (E-L stage 33), véraison (E-L stage 34), post-véraison (E-L stage 36) and ripe (E-L stage 38) (Coombe, 1995) in three sequential years (vintages: 2010/2011, 2011/2012 and 2012/2013). Berry samples (n=48 berries per sample) were collected between 10h00-11h00 on sampling days (Supplementary Table 3.1). The samples were flash frozen in liquid nitrogen in the field. Seeds were removed from the berries and the tissue homogenised in liquid nitrogen using an oscillating MM 400 mill (Retsch GmbH, Germany) and if not analysed immediately, stored at -80 °C until analyses.

Fresh samples were collected in the 2012/2013 season from each vine in the same two rows described above at green, pre-véraison, véraison and post-véraison stages. Ripe samples (E-L 38) could not be collected due as the vineyard harvest date was brought forward at the last moment and it was not possible to be in the field on time. Fresh samples were collected in perforated Falcon tubes and maintained on ice pips were removed before the berries were crushed into slurry with a mortar and pestle. Spectral analysis was carried out on the same day.

3.3.2 Extraction and HPLC analysis of sugars and organic acids

Sugars (glucose and fructose) and organic acids (malic acid, tartaric acid, and succinic acid) were extracted from 80-100 mg of frozen, ground berry tissue as described in Eyéghè-Bickong and co-workers (2012). Metabolite separation and analysis was carried out using an Agilent 1100 series HPLC system (Agilent Technologies©, California) equipped with a diode array detector (DAD) coupled to a refractive index detector (RID). The separation was performed on an Aminex HPX-87H ion exchange column (300 mm x 7.78 mm) protected with a Bio-Rad guard cartridge (30 mm x 4.6 mm) using 5 mM sulfuric acid as a mobile phase. The flow rate was 0.5 mL/min and the column maintained at 55℃. ChemStation Rev. A.10.02 software (Agilent Technologies©) was used for data acquisition and the injection volume was 10 µL. The DAD was used to detect organic acids (210 nm) while the RID was used to detect sugars. The sugars and organic acids were quantified using external standard calibration based on standard curves constructed using the peak areas versus the concentration (g/L) of authentic standards. Instrument detection limits for the different metabolites are shown on Supplementary Table 2. The standard concentrations ranged from 0.04 to 5 g/L for organic acids and 0.1 to 20 g/L for

sugars. The resulting concentrations were normalised to the internal standard amount and divided by the sample weight (fresh weight, FW) to obtain the concentration in mg/g FW. One-way ANOVA was performed with Statistica 12 software (StatSoft Inc., Tulsa, USA) on the reference data to determine the significance of the changes in metabolite levels between the developmental stages.

3.3.3 Infrared spectroscopy of homogenized berry samples

Sample preparation for IR analysis was adapted from a method described by Cozzolino et al. (2008). The IR spectra were generated from homogenates of frozen and fresh grape berry samples in reflectance mode. FT-NIR spectra were generated from the 2010/11 frozen samples using a Multi-Purpose Analyzer (MPA) FT-NIR spectrometer (Bruker Optics, Ettlingen, Germany), fitted with an integrating rotating sphere. A minimum of 5 g sample for each frozen grape berry homogenate was transferred from the freezer to a 50 ml Falcon tube and allowed to thaw at room temperature (±23 ^{o}C). Approximately 2.5 g was transferred to the spectrometer sample cup with a quartz base (diameter 3 cm). Samples were scanned against air background spectra that were measured at regular intervals during sample spectra acquisition under the same instrumental and laboratory conditions. After each measurement, the scanned sample was returned to the original Falcon tube, the contents mixed and another 2.5 g portion sampled and scanned. This process was repeated till ten technical repeats were acquired (600 spectral measurements). The sample cup was cleaned with distilled water and dried with lens-cleaning tissue paper between measurements. Sample measurement time was 45 seconds using the following scanning parameters: 12000–4000 cm^{-1} wavenumber range, 4 cm^{-1} resolution, 10 kHz scanner velocity, with 32 background and samples scans. The instrumental control and initial spectral processing were carried out using OPUS software (OPUS v. 7.0 for Microsoft, Bruker Optics, Ettlingen, Germany). NIR spectra were corrected for baseline offset in OPUS and a mean sample spectrum was acquired from the respective technical repeats.

ATR FT-MIR spectra were acquired at the same time as the NIR spectra. Approximately 45 mg of the same homogenate thawed for NIR analysis was scanned on an Alpha-P ATR FT-MIR spectrometer (Bruker Optics, Ettlingen, Germany), fitted with a single bounce diamond crystal sample plate (area 2 mm^2) maintained at 30^{o}C. After each measurement, the scanned material was discarded, the sample plate cleaned and a new sample scanned until ten technical repeats were acquired (1200

spectral measurements). MIR analysis was done under the following scanning parameters: 4000–400 cm^{-1} wavenumber range, 4 cm^{-1} resolution, 7.5 kHz scanner velocity, and 64 background and sample scans. Spectra were collected against a background of air as previously described. From each set of ten independent technical repeats; an average sample spectrum was generated in OPUS as previously described. No spectral pre-processing was applied.

In order to assess the suitability of ATR FT-MIR spectroscopy for use on fresh grape berries, homogenates of 192 fresh grape samples from the 2012/13 season were used. Due to the high repeatability and reproducibility observed with the frozen berry ATR FT-MIR spectra, the number of technical repeats on fresh samples was reduced to three (576 spectral measurements) and the average spectra of each sample was generated as previously described.

3.3.4 Multivariate data analysis of spectral data

OPUS was used for spectral acquisition; data processing was performed using SIMCA 13.0 (Umetrics, Umeå, Sweden). Spectral filtering was done using standard normal variate (SNV) correction, Savitzky−Golay transformation (first derivative and second derivatives, 15 points) and multiplicative scattering correction (MSC) were applied on the averaged baseline corrected FT-NIR sample spectra. These methods have been described to remove unwanted variation in spectral data (Blanco & Villaroya, 2002; Shenk et al., 2008; Shiroma & Rodriguez-Saona, 2009). Qualitative analysis (modelling of developmental stage) was carried out using PCA and OPLS-DA and quantitative analysis (calibration models and prediction of metabolites) was carried out with PLS. The spectral data was used as the X matrix and the reference data served as the Y matrix. For model development, the X block was mean-centered for PCA and PLS and Pareto-scaled for OPLS-DA and the Y block was scaled to unit variance.

3.3.4.1 Unsupervised clustering (PCA)

PCA in X was carried out to explore the possible clustering of samples and evaluate the influence of developmental stage. PCA is an unsupervised data compression technique that maximises the variation in X data and projects the main variation onto a few latent variables (Wold et al., 1987). Sample groupings, if present, are evident as clusters in PCA scores plots with the corresponding loadings plots showing the variables that contribute most to sample groupings (Wold et al., 1987).

3.3.4.2 Supervised clustering/discriminant analysis (OPLS-DA)

OPLS-DA, a supervised classification technique that can isolate the predictive component from the orthogonal sources of variation in the datasets and the S-line plot which identifies the discriminant variables responsible for the between-stage and within stage demarcations (Wiklund et al., 2008) were used to examine if sample classification could be improved in cases where PCA could not show clear clustering. OPLS-DA works through projection of X and is guided by known class information thus offering increased separation projection in comparison to PCA (Trygg et al., 2007; Wilkund et al., 2008). OPLS-DA models are easier to interpret as their loadings plots are not subject to orthogonal variation (Wilkund et al., 2008). A dummy variable was used to designate stage membership with a value of 1 assigned to samples that belonged to a specific stage and a value of 0 to samples that belonged to the four other stages under investigation. Spectral data was Pareto-scaled and the dummy variables were scaled to unit variance for OPLS-DA. The S-line plot combines the modelled covariance and correlation in an OPLS-DA model and gives a visual image of the OPLS-DA loadings with high modelling and discriminatory power (Wiklund et al., 2008). The OPLS-DA models were constructed between each two successive berry development stages.

3.3.4.3 Regression analysis of spectral data (PLS)

PLS regression analysis was carried out to construct calibration models for predicting the concentrations of organic acids (tartaric acid, malic acid, succinic acid), and sugars (glucose and fructose) in berry samples. PLS relates two data matrices X (wavenumbers) and Y (HPLC reference data) by a bilinear multivariate relationship (Wold et al., 2001). PLS regression analysis is a powerful tool for handling data with missing data in both X and Y, noise and collinearity between the variables, such as seen with spectroscopic data (Wold et al., 2001). The samples for the calibration and test sets were selected using multivariate sampling from PCA scores plots of MIR spectra and NIR spectra to get near equal representative samples from the four quadrants.

The 1st derivative Savitzky-Golay NIR spectra from the 2010/11 vintage were used to construct NIR calibration models. Due to the relatively small NIR sample size (n = 60), the calibrations were subjected to segmented cross-validation (seven samples per segment). The MIR calibration models were constructed using raw spectra of the

combined 2010/11 and 2011/12 seasons (n = 105), using the 1495-947 cm^{-1} range as described in Shiroma & Rodriguez-Saona (2009). Calibration and test sets were selected on a 2:1 split using multivariate sampling as previously described. The MIR calibration sets consisted of 73 samples and the test sets of 32 samples.

The statistics to evaluate the calibration set were the root mean square error of cross validation (RMSECV) to evaluate error in calibration, the root mean square error of estimation (RMSEE) to test the predictive accuracy of the calibration models and the coefficient of determination for calibration (R^2) to evaluate the fit of the models, the root mean square error of prediction (RMSEP) was used in the cases of test set validation. The bias (systematic difference between the estimated and true values) and the ratio of prediction to deviation were used to evaluate the predictive ability of the models. Calibration models that gave RPD values > 2 were considered good for screening, whilst RPD >3 was considered good for prediction, according to the interpretation by Versari et al. (2008).

3.4 Results and discussion

The overall aim of this study was to evaluate the suitability of FT-NIR and ATR FT-MIR spectroscopy to assess grape berry development qualitatively and quantitatively. Within this framework the objective was to develop simple, accurate (reliable/reproducible and preferably economical) protocols to determine quality-associated metabolites in grape berry samples.

3.4.1 Grape composition: HPLC analysis

A summary of the reference data for the different berry developmental stages as determined by HPLC is provided in Table 3.1. The concentration of tartaric acid declined throughout the whole developmental cycle, from green (13.5 mg/g FW) to pre-véraison (12 mg/g FW), to véraison (10.1 mg/g FW) and plateaued at the post-véraison stage (6.6 mg/g FW). The malic acid concentration doubled from green (7.7 mg/g FW) to pre-véraison (14.1 mg/g FW), before decreasing from véraison to ripe (Table 1). Concentrations of succinic acid fluctuated within a narrow range (2.9 – 4.3 mg/g FW) between green and post-véraison samples and decreased at the ripe stage (1.3 mg/g FW). Fructose concentration was low at the green and pre-véraison stages (0.7 – 2.3 mg/g FW) and increased at véraison (31.3 mg/g FW) before stabilising at the post-véraison and ripe stages (64.3 – 78.5 mg/g FW). The trend for glucose was similar to that of fructose. The observed sugar and organic acid profiles were consistent with literature reports at green, véraison, post-véraison and ripe stages (Hunter et al., 1991; Coombe 1992; Ali et al., 2011; Muñoz-Robredo et al., 2011).

Table 3.1: Reference data for organic acids and sugars (mg/g FW) showing the concentration range (minimum – maximum), mean and standard deviation (SD) at different stages of grape berry development. Non-identical alphabet letters (*a-d*) in the same row indicate significant differences in metabolite concentration between stages at $p < 0.01$.

Metabolite	STAGE									
	Green (n=21)		Pre-véraison (n=21)		Véraison (n=21)		Post-véraison (n=21)		Ripe (n=21)	
	Min-Max	Mean±SD	Min-Max	Mean±SD	Min-Max	Mean±SD	Min-Max	Mean±SD	Min-Max	Mean±SD
Malic acid	4.7-10.2	7.7±1.9a	8.9-16.6	14.1±1.8b	8.0-17.4	11.6±2.3c	1.3-14.9	3.5±2.8d	1.3-4.9	2.5±0.9d
Succinic acid	2.2-7.7	4.3±1.6a	1.6-5.0	2.9±1.5b	1.5-6.0	4.0±1.4c	0.9-3.9	2.3±1.5d	0.6-3.0	1.3±1.5d
Tartaric acid	11.6-16.3	13.5±1.3a	8.1-15.6	12.0±1.5a	8.3-12.4	10.1±1.1b	4.7-11.8	6.6±1.5c	5.4-7.8	6.6±0.7c
Fructose	0.2-1.6	0.7±0.4a	1.0-8.4	2.3±1.8a	5.8-44.8	31.3±10.5b	54.5 -75.7	64.3±15.0c	58.0-82.5	73.5±5.6d
Glucose	1.6-3.7	2.9±0.5a	2.9-14.9	5.9±2.9a	11.3-50.0	37.1±10.5b	55.8-77.5	65.5±14.6c	58.1-82.2	73.5±5.3d

3.4.2 Description of the IR (FT-NIR and ATR FT-MIR) spectral features from grape berry samples

For each developmental stage, a mean baseline corrected NIR spectrum of the frozen samples was generated and the spectra at the five stages are shown in Fig. 3.2 A. NIR spectra did not show obvious systematic differences that could be associated with the respective berry development stages. The NIR spectrum consists of overtones and combination bands of the fundamental vibrations observed in the MIR region (Shenk et al., 2008). Prominent NIR absorption peaks were observed at 10217, 8436, 6978, 5643, and 5614 cm^{-1} (Fig. 3.2 A).

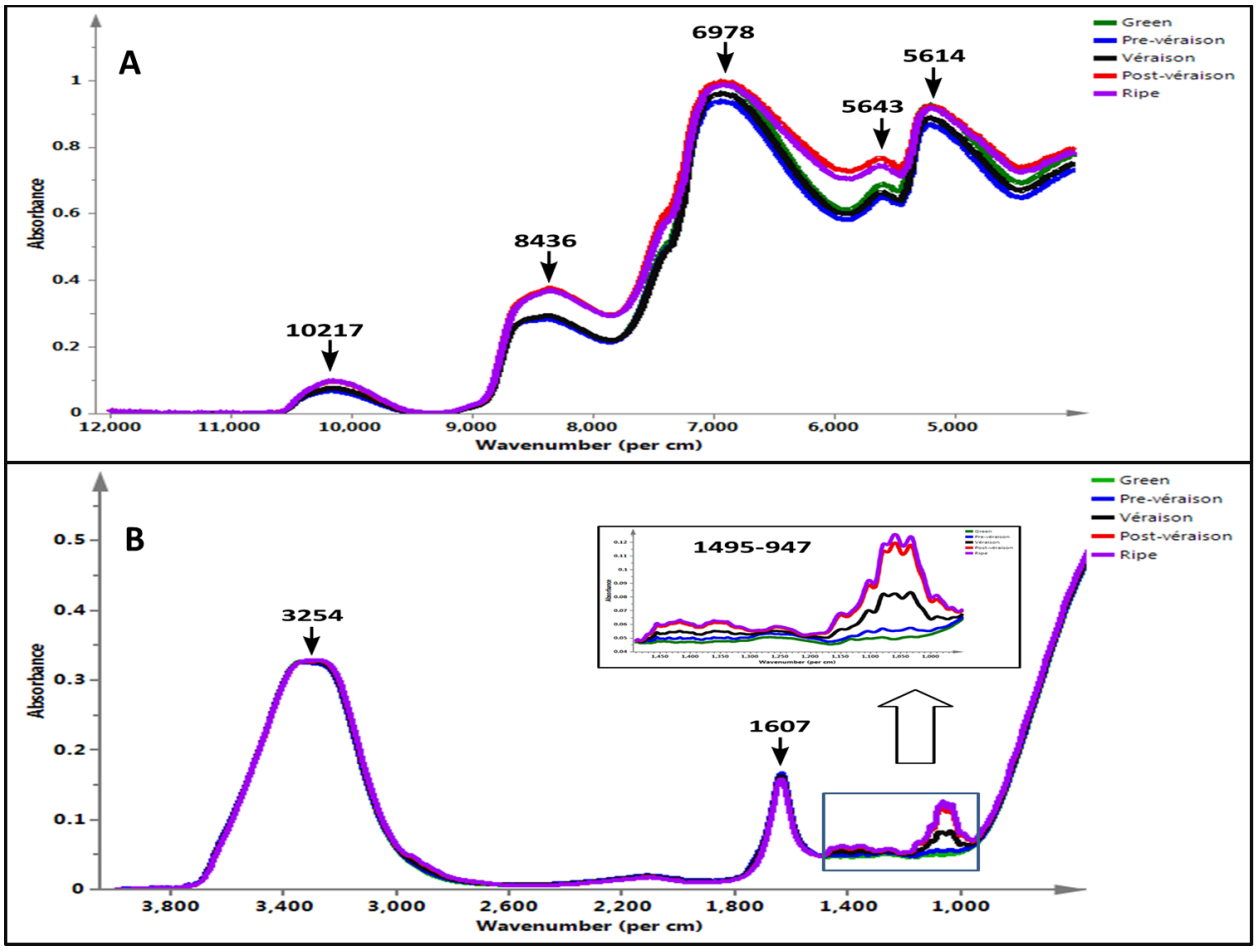


Fig 3.2. NIR spectra (Fig. 3.2A) and MIR spectra (Fig. 3.2B) of frozen grape berries at five developmental stages (as coloured in legend). NIR peak annotations correspond to the absorptions of water (10217, 8436, 6978 and 5614 cm^{-1}), sugars and organic acids (6978, 5643 and 5614 cm^{-1}); MIR peak annotations correspond to the absorptions of water (3254 and 1607 cm^{-1}) and fingerprint region (insert, 1495-947 cm^{-1}).

Peak assignment was done in accordance with the literature (Nicolaï et al, 2007; Shenk et al., 2008; Shiroma & Rodriguez-Saona 2009). Whereas many compounds

absorb in the NIR bands highlighted here, contributions of the O–H overtone or combination in water were observed at 10217, 6978, 5643 and 5614 cm^{-1}. The C–H overtone and stretch in sugars and organic acids were observed at 8436 and 6978 cm^{-1} whilst the C–O stretch and overtone in sugars and organic acids were ascribed at 6978, 5643 and 5614 cm^{-1}.

MIR spectra for both fresh and frozen berries were highly reproducible and had a distinct region of systematic variation related to developmental stage observed at 1495-947 cm^{-1} (Fig. 3.2B). The region of systematic variation (Fig. 3.2B, insert) was called the fingerprint region as previously reported by Shiroma & Rodriguez-Saona, 2009 and is known to be associated with C-O and C-C stretches in sugars, esters and organic acids amongst other compounds. Further band assignment was done according to literature as follows: 3300-3200 and 1710-1590 cm^{-1} corresponding to the O–H stretch in water with prominent peaks observed at 3254 and 1607 cm^{-1}; 1480-1390, 1420-1320 and 1150-1060 cm^{-1} ascribed to C–H, O–H and C–O functional groups respectively in organic acids; 1100-950 cm^{-1} due to the contribution of the C–O stretch related to carbohydrates with the most intense bands located at 1061 and 1033 cm^{-1} emanating from fructose and glucose respectively (Kelly et al., 2004; Moreira & Santos, 2004; Shiroma & Rodriguez-Saona, 2009). The MIR spectra for fresh grape berries were visually very similar to the frozen samples when assessed at the same stage of development (data not shown).

3.4.3 Multivariate data analysis of IR spectral data

3.4.3.1 PCA

For preliminary examination of both NIR and MIR spectra, PCA was performed to examine time-related changes that occurred during grape berry development. Scores plots from baseline corrected NIR spectra revealed no clear groupings according to developmental stage (Fig. 3.3A). Application of spectral filters Savitzky-Golay 1st derivative, SNV, MSC and OSC did not improve the stage discrimination as illustrated by the PCA scores plot of the Savitzky-Golay 1st derivative spectra in Fig. 3.3B.

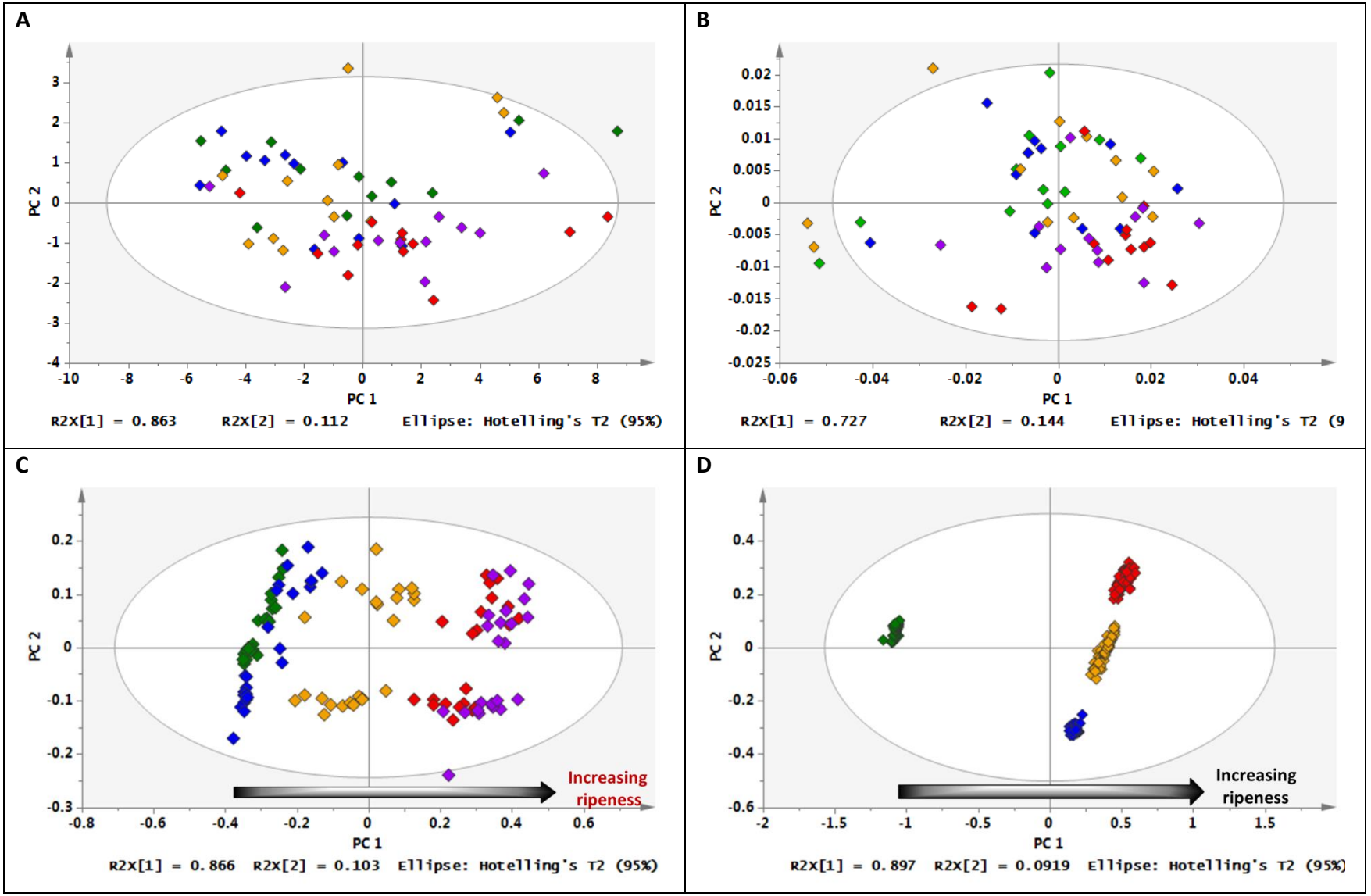


Fig 3.3: PCA scores plots NIR (A); NIR Savitzky-Golay 1st derivative (B) MIR (frozen berries, C) and MIR (fresh berries, D). The colours represent different developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).

PCA scores plots obtained from ATR FT-MIR spectra of frozen samples (Fig. 3.3C) showed a discernible pattern of increasing berry ripeness along PC 1. Pre-véraison samples co-clustered with green samples and post-véraison samples co-clustered with ripe samples. The samples on Fig 3.3 C appeared to be separating according to vintage across PC 2 with the 2010/11 vintage on the left side of PC 2 1nd the 2011/12 vintage on the right; this could not be ascertained as two machines were used for the different seasons. Examination of the PCA scores plot generated from fresh berries showed very well-defined sample clusters for each of the four stages tested as similar clustering patterns to those observed with frozen berries were observed on fresh berries with increasing ripeness across PC 1 (Fig. 3.3D).

3.4.3.2 OPLS-DA

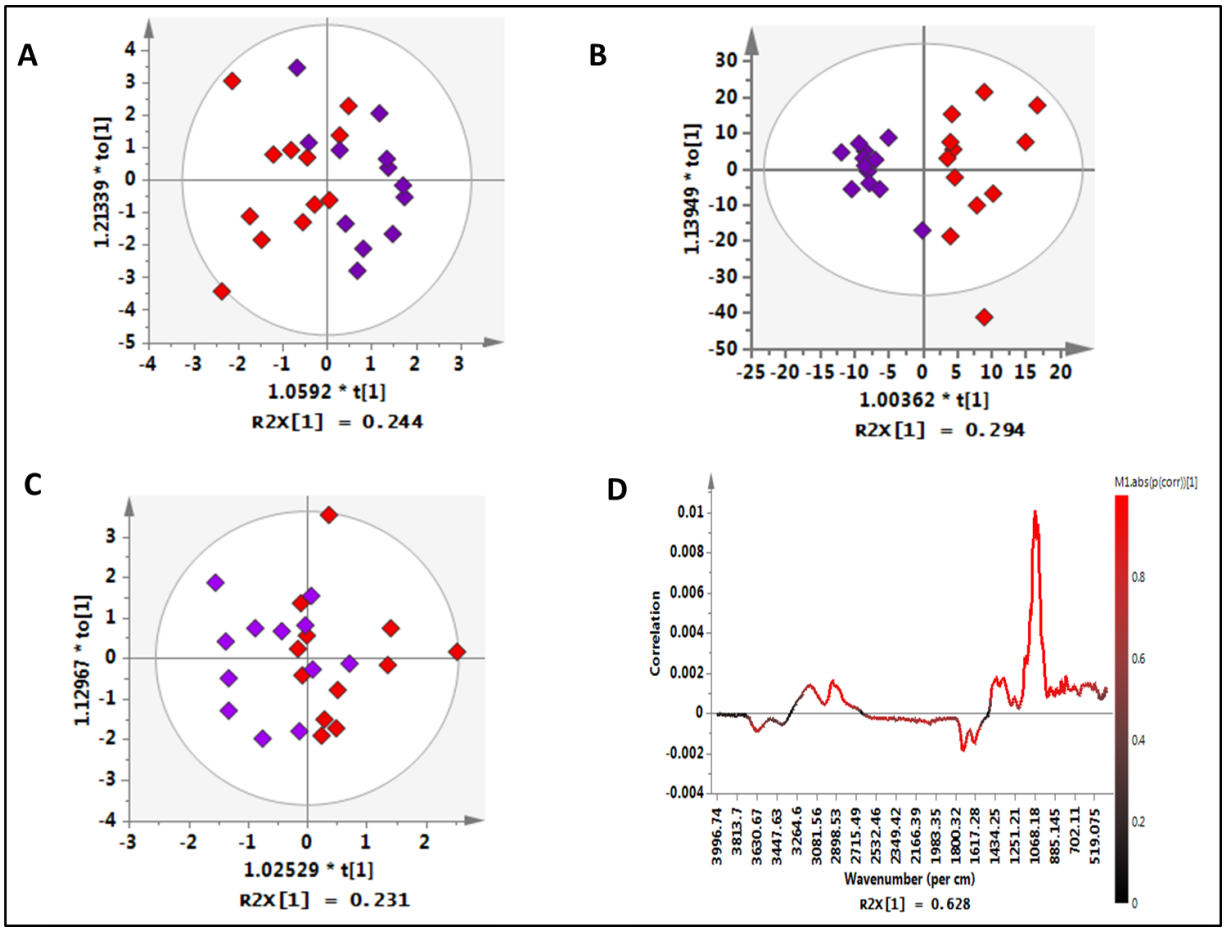


Fig 3.4: OPLS-DA scores plots for post-véraison (◆)/ripe (◆) with HPLC reference data (A), NIR baseline corrected spectra (B), and MIR spectra (C). The S-line plot (véraison/post véraison)-(D) showed the wavenumbers responsible for between stage discrimination (1104–1032 cm^{-1}) in the MIR region

Using OPLS-DA, it was possible to discriminate between every two successive stages of berry development with NIR spectral data Fig 3.4 A – C shows the OPLS-DA scores plots for reference data and spectral data at post-véraison /ripe stages as derived from frozen berry samples. Like with NIR data, OPLS-DA distinguished all but post-véraison and ripe stages with MIR and reference data (Supplementary Fig. 3.1). This observation showed that despite its simplified approach, IR spectroscopy is a powerful tool that can provide results than can be used to define berry growth and development without extensive sample preparation.

Despite similar ripening patterns being observed on fresh and frozen samples by MIR, fresh berries showed more discrete clusters. Cold storage potentially causes structural and chemical changes in grapes as they undergo additional chemical changes from frozen to thawed states (Cozzolino et al., 2005). The purported structural and chemical distortions taking place between freezing and thawing were not apparent on the spectra and the nature or extent of such chemical changes was beyond the scope of this study.

The ability to successfully discriminate the different developmental stages by both PCA and OPLS-DA using IR spectra could be taken as indicators that there are significant differences between developmental stages which FT-NIR and ATR FT-MIR spectroscopy are both capable of elucidating.

Furthermore from OPLS-DA, the S-line plot showed the MIR wavenumbers responsible for stage discriminatory and narrowed them to 1104–1032 cm^{-1}(Fig. 4C). These were the same wavenumbers associated with absorbance of molecules such as the C–O, O–H and C–C bond stretches of phenolics, organic acids and sugars as interpreted by Moreira & Santos (2005) Shiroma & Rodriguez-Saona (2009).

3.4.3.4 PLS regression models

The best NIR calibration models were developed using the 15 point Savitzky-Golay 1^{st} derivative. NIR spectra were used for developing preliminary models only and additional work would be carried out to get more robust models. For MIR calibrations, the best models were constructed from the fingerprint region. The MIR calibration models for tartaric acid, malic acid, glucose and fructose are shown in Fig. 3.5.

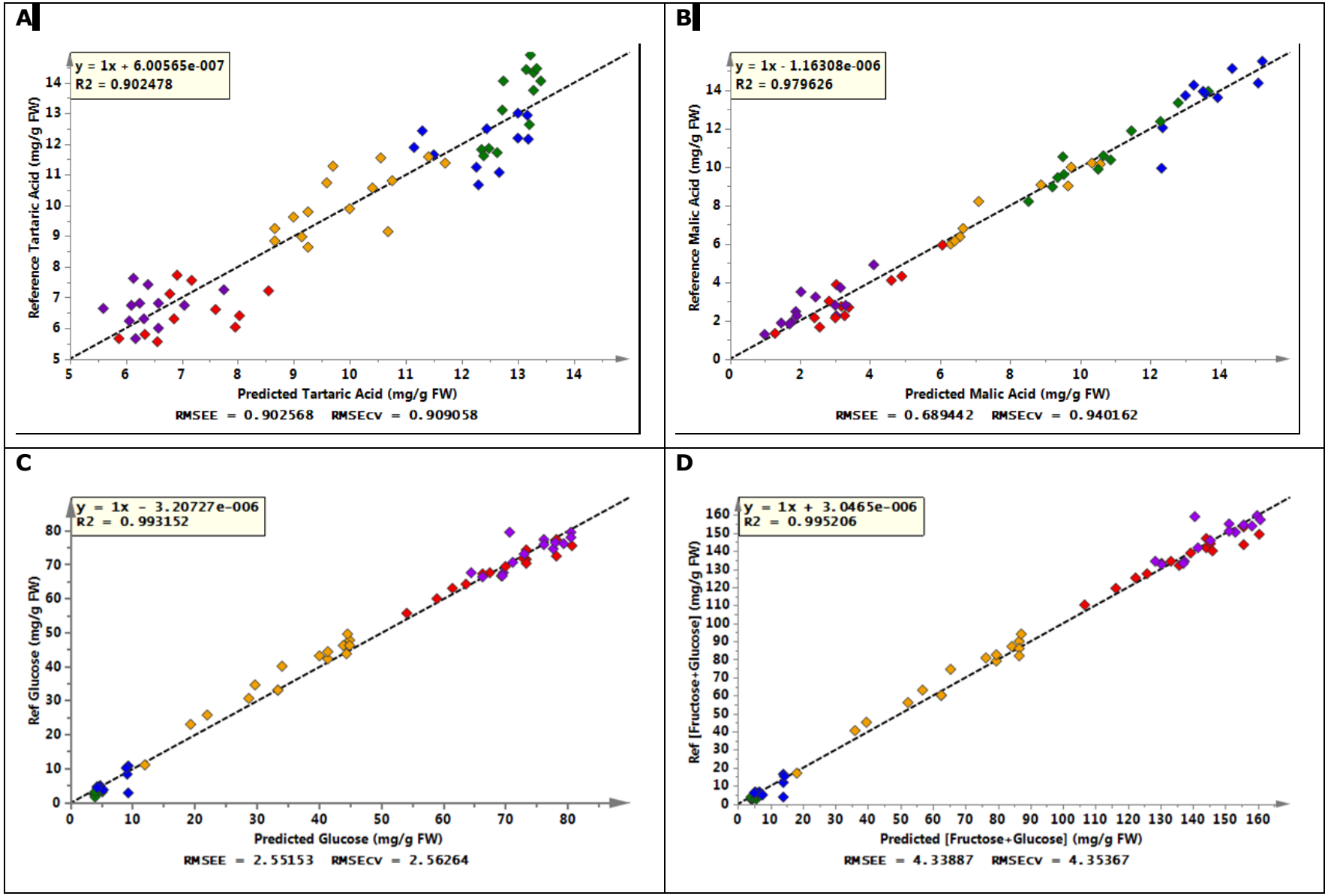


Fig. 3.5: PLS regression models from MIR spectra for tartaric acid (A), malic acid (B), glucose (C), glucose + fructose (D) showing the development stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆)

The statistical indicators for model fitness showed that both NIR and MIR spectra yielded relatively accurate calibration models. When compared to NIR models, the MIR models were more accurate (Table 3.2).

Table 3.2: Model evaluation statistics for calibration (R^2, RMSECV and RMSEE) in white and test sets (Bias, RMSEP and RPD, shaded) as determined from FT-NIR and ATR FT-MIR spectra.

Analyte	Set	Evaluation statistics	FT-NIR	ATR FT-MIR
Tartaric acid	Calibration	[a]R^2	0.81	0.9
		[b]RMSECV (mg/g)	1.6	0.9
		[c]RMSEE (mg/g)	1.4	0.9
	Test	[d]Bias	-0.33	-0.21
		[e]RMSEP (mg/g)	3.0	0.9
		[f]RPD	1.0	3.3
Malic acid	Calibration	R^2	0.69	0.97
		RMSECV (mg/g)	3.1	0.7
		RMSEE (mg/g)	2.9	0.9
	Test	Bias	-	-0.02
		RMSEP (mg/g)	-	1.1
		RPD	-	2.9
Succinic acid	Calibration	R^2	0.56	0.76
		RMSECV (mg/g)	0.8	1.1
		RMSEE (mg/g)	0.8	0.9
	Test	Bias	-	-0.02
		RMSEP (mg/g)	-	1.1
		RPD	-	1.7
Glucose	Calibration	R^2	0.83	0.99
		RMSECV (mg/g)	13.1	2.6
		RMSEE (mg/g)	12.8	2.6
	Test	Bias	-	0.13
		RMSEP (mg/g)	-	5.3
		RPD	-	10.5
Fructose	Calibration	R^2	0.85	0.99
		RMSECV (mg/g)	13.2	3.2
		RMSEE (mg/g)	12.8	3.2
	Test	Bias	-	-0.03
		RMSEP (mg/g)	-	3.4
		RPD	-	7.9

[a]R^2: coefficient of determination

[b]RMSECV: root mean square error of cross validation

[c]RMSEE: root mean square error of estimation

[d]RMSEP: root mean square error of prediction

[e]RPD: ratio of prediction to deviation

[f]Bias: systematic difference between the estimated and true values

R^2 values were high for organic acids (>0.81 for malic and tartaric acids) and sugars (> 0.98 for fructose and glucose). RMSEP was 1.2 mg/g FW for organic acids, and 4.7 mg/g FW for sugars. RPD values were > 3.0 for tartaric acid and > 7.0 for glucose and fructose (Table 2). FT-NIR has been used to evaluate ripening of Sangiovese grapes and the authors reported high coefficients of determination for sugars (R^2 = 0.92) and lower coefficients of determination for organic acids (R^2 =0.47) (Barnaba et al., 2013). Similarly, ATR FT-MIR has been used to determine phenolics and tannins in grape berries and high coefficients of determination (R^2 >0.95) with high RPDs (>2.9) reported (Fragoso et al., 2011).

The PLS model for fructose slightly deteriorated at green and pre-véraison stages. At these stages the fructose concentrations between 0.2 – 8.4 mg/g FW were generally below 0.3 mg/g FW at green and deviation exceeded 10% in some cases. The limit of quantitation provided by the reference method for fructose was 0.1 mg/g FW (RID) and 0.3 mg/g FW (DAD). FTIR spectroscopy is known to be insensitive to concentration levels below 200 ppm (0.2 mg/g) (Bauer et al, 2008). Thus at the green stage the deviations between the reference and predicted values were generally > 10% as the sensitivity of the method edged towards its limits. However the prediction was very good at véraison, post-véraison and ripe stages with sugar levels > 25 mg/g FW (deviation < 10%). To explore the possibility of generating better prediction models for sugars, sectional calibration models were developed for glucose and fructose at green and pre-véraison stages on the one hand, and, véraison, post-véraison stages on the other, but this did not improve the model prediction (data not shown). This observation showed that the model could not be relied on to predict fructose at concentrations below 0.3 mg/g FW. However when glucose and fructose were calibrated together the prediction improved significantly at green and pre-véraison stages (prediction error < 10%) and from véraison to ripe the prediction error dropped to less than 7%. The sugar concentrations at post-véraison and ripe matched the values reported in the literature. Barnaba et al. (2013) reported 105.9 and 113.0 mg/g FW for ripe Sangiovese grapes and González-Caballero et al. (2011) reported 221.7 mg/g for reducing sugars in ripe berries from white and red cultivars whilst Fernández-Novales et al. (2009) reported fructose concentrations of 136.1 mg/g in ripe red/white grapes.

For malic and tartaric acids, prediction was quite good across the entire growth curve with prediction error from the reference data hardly ever exceeding 10%. The concentrations of tartaric and malic acids at post-véraison ripe grapes were in the

same ranges as reported by González-Caballero et al., (2011) when they assessed the quality of ripening and ripe white and red grapes (tartaric acid: 6.0–12.2 mg/g, malic acid: 0.3–2.0 mg/g). Barnaba et al., (2013) reported similar organic acid concentrations in ripening Sangiovese grapes (tartaric acid: 7.3 mg/g, tartaric acid: 1.5 mg/g). Succinic acid (RPD < 1.6 and R^2 < 0.75) though generally well predicted had the widest deviation range (6 – 32%). The fact that a relatively low abundance metabolite like succinic acid could be accurately predicted in over 70% of the cases showed that the results obtained here offer a promising starting point for the development of quantitative assessment techniques that can be applicable to some of the less abundant grape berry metabolites.

MIR was shown to be better suited for both qualitative and quantitative applications. This is most likely due to the fact that the fundamental vibrational absorption bands in MIR are better resolved and show better specificity and reproducibility than the combination and overtone bands of the NIR (Kelly et al., 2004). An additional advantage of the ATR FT-MIR spectroscopic method is the temperature controlled ATR crystal that reduces potential variation by maintaining constant sample temperature (Smyth & Cozzolino, 2007).

3.4.3.5 Prediction of metabolite concentrations

The ultimate objective for constructing PLS regression models was to come up with models that could be used to predict sugars and organic acid concentrations in grape berry samples during the entire growth cycle (Supplementary Table 3.2). It was observed that the predicted values for both the frozen and fresh samples were very close to those presented by the reference data with general deviation between reference and predicted values well below 10%. This was regardless of the different residence time of the samples at the laboratory prior to analysis fresh. Interestingly even succinic acid, which had the lowest R^2 and RPD values, was well predicted for most of the samples.

To evaluate the robustness of the models, 14 independent frozen samples with known reference values were tested. The prediction error of the predicted concentrations versus the analytical reference values ranged between 1.3 – 8.4% for all metabolites. Using the predicted results from the models it was observed that organic acid concentrations were highest at green and pre-véraison stages and gradually declined to a steady-state after véraison. Sugar concentrations were lowest at green and pre-véraison stages, but exponentially increased till the véraison stage. Sugar

concentrations typically plateaued at the post-véraison stage remaining relatively constant till the ripe stage (Fig. 3.6).

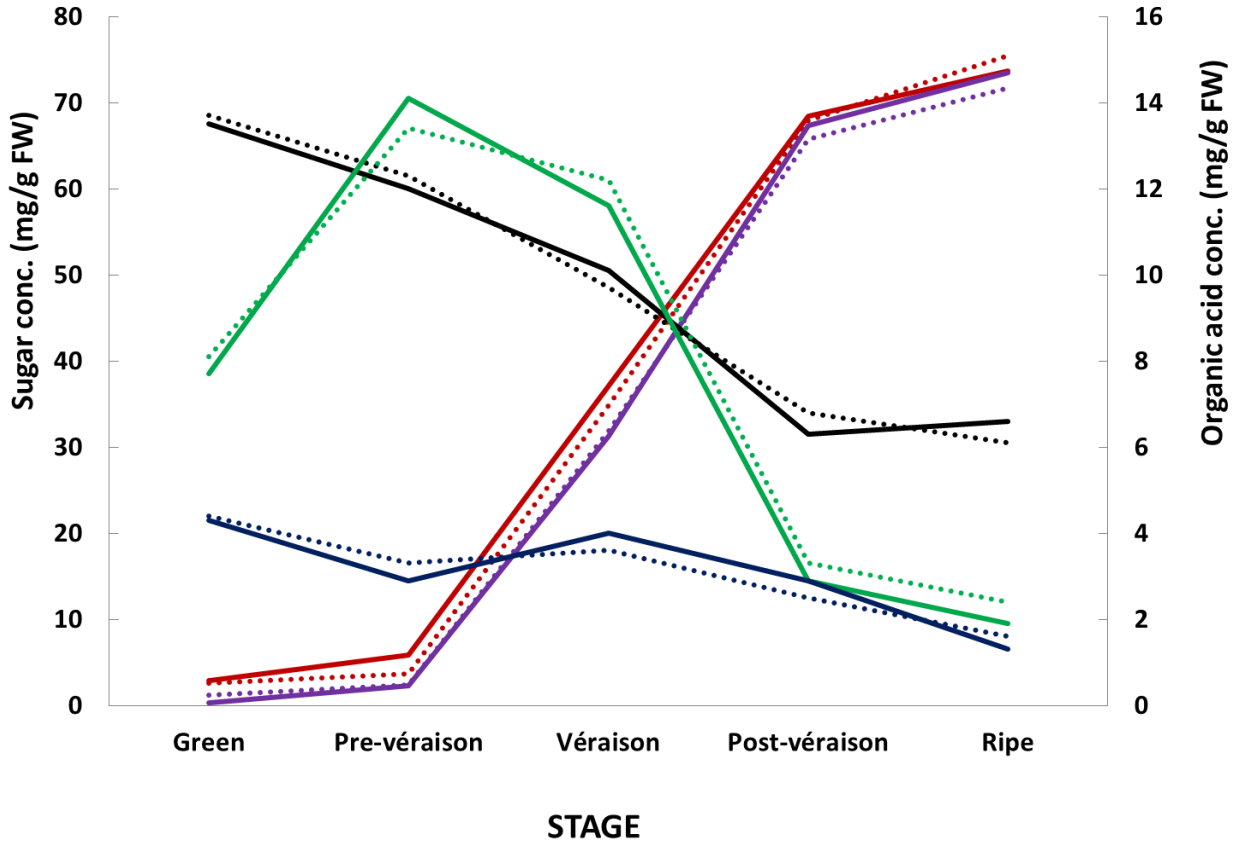


Fig 3.6: Average grape berry metabolite concentration changes predicted by MIR spectroscopy over five developmental stages: tartaric acid (◆), malic acid (◆), succinic acid (◆), glucose (◆) and fructose (◆).The continuous curves represent reference data graphs and the dotted graphs represent curves constructed from predicted data.

The predicted metabolite profiles were in good accordance with the reported developmental profiles/trends of these metabolites in grape berries (Topalovic & Mikulic-Petkovsek, 2010; Muñoz-Robredo et al., 2011). A diagrammatic depiction of grape berry metabolite growth curve based on averaged predicted frozen values is shown in Fig. 3.6.

3.5 Conclusions

Classification of grape berries according to their respective developmental stage was possible using ATR FT-MIR. Interestingly the ATR FT-MIR spectra from fresh homogenised grape berry samples resulted in better developmental stage discrimination (by stage comparison) than the spectra from frozen homogenized samples using PCA, and the quantitative analytical reference data (HPLC). Although overall developmental progression between the respective stages were observable with MIR spectra (using PCA and OPLS-DA) and NIR spectra (using OPLS-DA); stage discrimination at the post-véraison/ripe stages was only possible with NIR spectra. Needless to say, the difficulty in separating the latter two stages highlights the rather obvious fact that changes during grape berry development involves more than a few selected compounds (e.g. sugars and organic acids). Non-targeted analytical methods (*e.g.* IR spectrocopy) provide additional information for stage-specific discrimination. The identification of the compounds underlying this discrimination in the closely related stages will provide interesting information on berry development. Results for the OPLS-DA and S-line plot showed that the sugar/organic acid absorption band was key to between-stage variation.

Grape berry composition is an objective indicator of the largely subjective concept of grape quality. The IR and MVDA technique described provide a fast method for the relatively high throughput qualitative and quantitative analysis of berry samples, without the need for metabolite extraction/purification. Both FT-NIR and ATR FT-MIR spectroscopy are demonstrated to be versatile techniques for qualitative and quantitative determination of the sugar and organic acids in grape berries. The approaches provide a powerful way to rapidly extract qualitative and quantitative information emanating from multiple spectral variables.

3.6 Acknowledgements

Thanks are due to Professor Martin Kidd (Department of Statistics and Actuarial Sciences, Stellenbosch University, South Africa) for his assistance with statistical analysis. The advice rendered by Professor Johan Trygg (Department of Chemistry, Umeå University, Sweden) is gratefully acknowledged. The authors thank the financial support provided by the South African Wine industry (Winetech) as well as the Technology and Human Resource for Industry Programme of South Africa.

3.7 Competing interests

There are no competing interests.

3.8 Appendix

Supplementary tables and figures

Supplementary Table 3.1: Sampling dates in the three consecutive years

Supplementary Figure 3.1A: OPLS-DA scores plots for reference values as analysed at green/pre-véraison (A), pre-véraison/véraison (B), véraison/ post-véraison (C), and post-véraison/ripe (D). The colours represent the developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).

Supplementary Figure 3.1B:. OPLS-DA scores plots NIR spectra as analysed at green/pre-véraison (A), pre-véraison/véraison (B), véraison/ post-véraison (C), and post-véraison/ripe (D). The colours represent the developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).

Supplementary Figure 3.1C: OPLS-DA score plots and MIR spectra as analysed at green/pre-véraison (A), pre-véraison/véraison (B), véraison/post-véraison (C), and post-véraison/ripe (D). The colours represent the developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).

Supplementary Table 3.2: Average metabolite concentrations (mg/g FW) as determined by HPLC (i.e. reference data from frozen grapes), MIR (from frozen grapes) and MIR (from fresh berries).

3.9 References

Alexandersson, E., Jacobson, D., Vivier, M.A., Weckwerth, W., & Andreasson, E. (2014). Field-omics – understanding large-scale molecular data from field crops. *Frontiers in Plant Science, 5,* 1-6.

Adato, R., & Altug, H. (2013). In-situ ultra-sensitive infrared absorption spectroscopy of biomolecule interactions in real time with plasmonic nanoantennas. *Nature Communications,* 1-14.

Ali, K., Maltese, F., Fortes, A.M., Pais, M.S., Choi, Y.H., & Verpoorte, R. (2011). Monitoring biochemical changes during grape berry development in Portuguese cultivars by NMR spectroscopy. *Food Chemistry, 124,* 1760-1769.

Barnaba, F.E., Bellincontro, A., & Mencarelli, F. (2014). Portable NIR-AOTF spectroscopy combined with winery FTIR spectroscopy for an easy, rapid, in-fieldmonitoring of Sangiovese grape quality. *Journal of Agricultural and Food Chemistry, 94,* 1071-1077.

Bauer, R., Nieuwoudt, H., Bauer, F.F., Kossmann, K., Koch, K.R., & Esbensen, K.H. (2008). FTIR spectroscopy for grape and wine analysis. *Analytical Chemistry, 1,* 1371-1378.

Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: a rapid-response analytical tool. *Trends in Analytical Chemistry, 21,* 240-250.

Boccard, J., & Rutledge, D.N. (2013). A consensus orthogonal partial least squares discriminant analysis (OPLS-DA) strategy for multiblock Omics data fusion. *Analytica Chimica Acta, 769,* 30-39.

Coombe, B.G. (1992). Honorary research lecture: Research on the development and ripening of the grape berry. *American Journal of Enology and Viticulture, 43,* 101-110.

Coombe, B.G. (1995). Adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research 1,* 100-110.

Coombe, B.G., & McCarthy, M.G. (2000). Dynamics of grape berry growth and physiology of ripening. *Australian Journal of Grape and Wine Research 6,* 131-135.

Cozzolino, D., Cynkar, W.U., Dambergs, R.G., Janik, L. & Gishen, M. (2005). Effect of both homogenisation and storage on the spectra of red grapes and on the measurement of total anthocyanins, total soluble solids and pH by visual near infrared spectroscopy. *Journal of Near Infrared Spectroscopy, 13,* 213-223.

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,* 7631-7636.

Dal Santo, S., Tornielli, G.B., Zenoni, S., Fasoli, F., Farina, L., Anesi, A. Guzzo, F., Massimo Delledonne, M. & Pezzotti, M., (2013). The plasticity of the grapevine berry transcriptome. *Genome Biology, 14,* 1-18.

Eyéghè-Bickong, H.A.; Alexandersson, E.O., Gouws, L.M.; Young, P.R., & Vivier, M.A. (2012). Optimization of an HPLC method for the simultaneous quantification of the major sugars and organic acids in grapevine berries. *Journal of Chromatography B, 885-886,* 43-49.

Fernández-Novales, J., López, M-I., Sánchez, M-T., Morales, J., & González-Caballero, V. (2009). Shortwave-near infrared spectroscopy for determination of reducing sugar content during grape ripening, winemaking, and aging of white and red wines. *Food Research International, 42*, 285-291.

Fragoso, S., Aceña, L., Guasch, J. Busto, O., & Mestres, M. (2011). Application of FT-MIR spectroscopy for fast control of red grape phenolic ripening. *Journal of Agricultural and Food Chemistry, 59,* 2175-2183.

García de Cortázar-Atauri, I., Brisson, N., & Gaudillere, J.P. (2009). Performance of several models for predicting budburst date of grapevine *(Vitis vinifera L.). International Journal of Biometeorology, 53,* 317-326.

Gholami, M., Hayasaka, Y., Coombe, B.G., Jackson, J.F., Robinson, S.P., & Williams, P.J. (1995). Biosynthesis of flavour compounds in Muscat Gordo Blanco grape berries. *Australian Journal of Grape and Wine Research, 1*, 19-24.

Giovenzana, V., Beghi, R., Malegori, C., & Civelli. R. (2014). Wavelength selection with a view to a simplified handheld optical system to estimate grape ripeness. *American Journal of Enology and Viticulture, 65,* 117-123.

Gishen, M.; Dambergs, R.G., & Cozzolino, D. (2005). Grape and wine analysis – enhancing the power of spectroscopy with chemometrics:A review of some applications in the Australian wine industry. *Australian Journal of Grape and Wine Research 11,* 296-305.

González-Caballero, V., Sánchez, M-T, López, M-I., & Pérez-Marín, D. (2010). First steps towards the development of a non-destructive technique for the quality control of wine grapes during on-vine ripening and on arrival at the winery*. Journal of Food Engineering, 101,* 158-165.

Herrera, J., Guesalaga, A. & Agosin, E. (2003). Shortwave near infrared spectroscopy for non-destructive determination of maturity of wine grapes. *Measurement Science and Technology, 14,* 689-697.

Hunter J.J., Visser, J.H., & De Villiers, O.T. (1991). Preparation of grapes and extraction of sugars and organic acids for determination by high performance liquid chromatography. *American Journal of Enology and Viticulture, 43,* 237-244.

Jackson, D.I., & Lombard, P.B. (1993). Environmental and management practices affecting grape composition and wine quality - a review quality. *American Journal of Enology and Viticulture, 44,* 409-430.

Jarén, C., Ortuño, C., Arazuri, S., Arana, J.I., & Salvadores, M.C. (2001). Sugar determination in grapes using NIR technology. *International Journal of Infrared and Millimeter Waves, 22,* 1521-1530.

Kelly, J.F.D., Downey, G., & Fourtier, V. (2004). Initial study of honey adulteration by sugar solutions using mid-infrared (MIR) spectroscopy and chemometrics. *Journal of Agricultural and Food Chemistry, 52,* 33-39.

Moreira, J. L. & Santos, L. (2004). Spectroscopic interferences in Fourier transform infrared wine analysis. *Analytica Chimica Acta, 513,* 263-268.

Muñoz-Robredo, P., Robledo, P., Manríquez, D., Rosa Molina, R. & Defilippi, B.G. (2011). Characterization of sugars and organic acids in commercial varieties of table grapes. *Chilean Journal of Agricultural Research, 71,* 452-458.

Martens, H., & Naes, T. (1989). *Multivariate Calibration.* Toronto: John Wiley & Sons (Chapter 2).

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I. & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46,* 99-118.

Quilitzsch, R., Baranska, M., Schulz, H., & E. Hoberg, E. (2005). Fast determination of carrot quality by spectroscopy methods in the UV-VIS, NIR and IR range. *Journal of Applied Botany and Food Quality 79,* 163-167.

Robinson, S.P., & Davies, C. (2000). Molecular biology of grape berry ripening. *Australian Journal of Grape and Wine Research, 6,* 175-188.

Rolle, L., Torchio, F., Giacosa, S., Río Segade, S., Cagnasso, E., & Gerbi, V. (2012). Assessment of physicochemical differences in Nebbiolo grape berries from different production areas and sorted by flotation. *American Journal of Enology and Viticulture, 63,* 195-204.

Rusjan, D., Korošec-Koruza, Z., & Veberič, R. (2008). Primary and secondary metabolites related to the quality potential of table grape varieties *(Vitis vinifera L.). European Journal of Horticultural Science, 73,* 124-130.

Shenk, J.S., Workman Jr., J.J., & Westerhaus, M.O. (2008). Application of NIR spectroscopy to agricultural products. In: Burns, D.A., & Ciurczik E.W. (Eds), *Handbook of NIR Analysis (pp.* 349-360). New York: CRC Press.

Shiroma, C., & Rodriguez-Saona, L. (2009). Application of NIR and MIR spectroscopy in quality control of potato chips. *Journal of Food Composition and Analysis, 22,* 596–605.

S̆uklje, K., Coetzee, Z., Antalick, G., Cesnik, H.B., Brandt, J., Lisjak, K., & Deloire, A. (2012). Classification of grape berries according to diameter and total soluble solids to study the effect of light and temperature on methoxypyrazine, glutathione, and hydroxycinnamate evolution during ripening of *Sauvignon blanc (Vitis vinifera L.)*. *Journal of Agricultural and Food Chemistry, 60,* 9454-9461.

Smyth, H.E., & Cozzolino, D. (2011). Applications of infrared spectroscopy for quantitative analysis of volatile and secondary metabolites in plant materials. *Current Bioactive Compounds, 7,* 66-74.

Swanepoel, M., Du Toit, M. & Nieuwoudt, H.H. (2007). Optimisation of the quantification of total soluble solids, pH and titratable acidity in South African grape must using Fourier transform mid-infrared spectroscopy. *South African Journal Of Enology And Viticulture, 28,* 140-149.

Topalovic, A., & Mikulic-Petkovsek, M. (2010). Changes in sugars, organic acids and phenolics of grape berries of cultivar Cardinal during ripening. *Journal of Food, Agriculture and Environment, 8,* 223-227.

Trygg, J., Holmes, E., & Lundstedt, T. (2007). Chemometrics in metabonomics. *Journal of Proteome Research 6,* 469-479.

Versari, A., Parpinello, G.P., Mattioli, A.U., & Galassi, S. (2008). Determination of grape quality at harvest using Fourier-transform mid-infrared spectroscopy and multivariate analysis. *American Journal of Enology and Viticulture 59,* 317-322.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E.J., Edlund, U., Shockcor, J.P., Gottfries, J., Moritz, T., & Trygg, J. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry, 80,* 115-122.

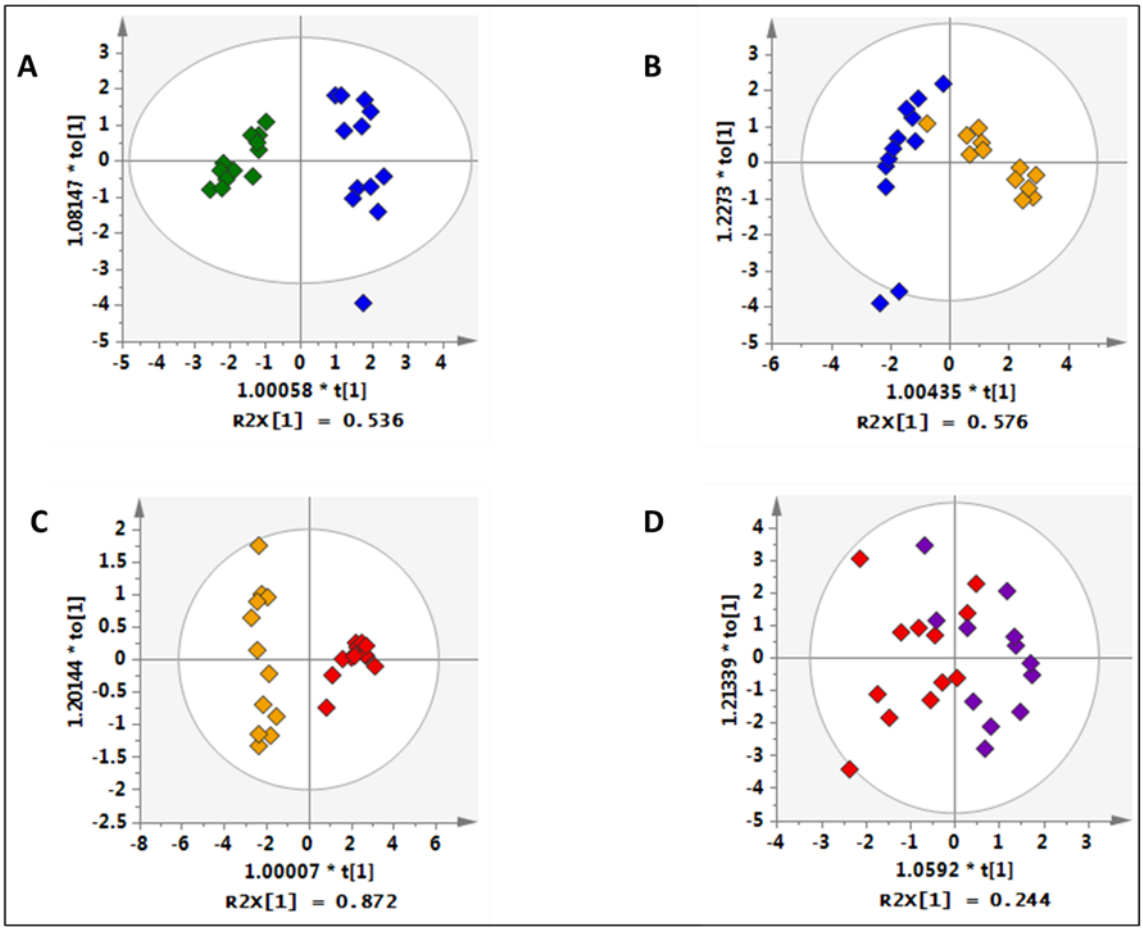
Wold S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2,* 37-52.

Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems, 58,* 109–130.

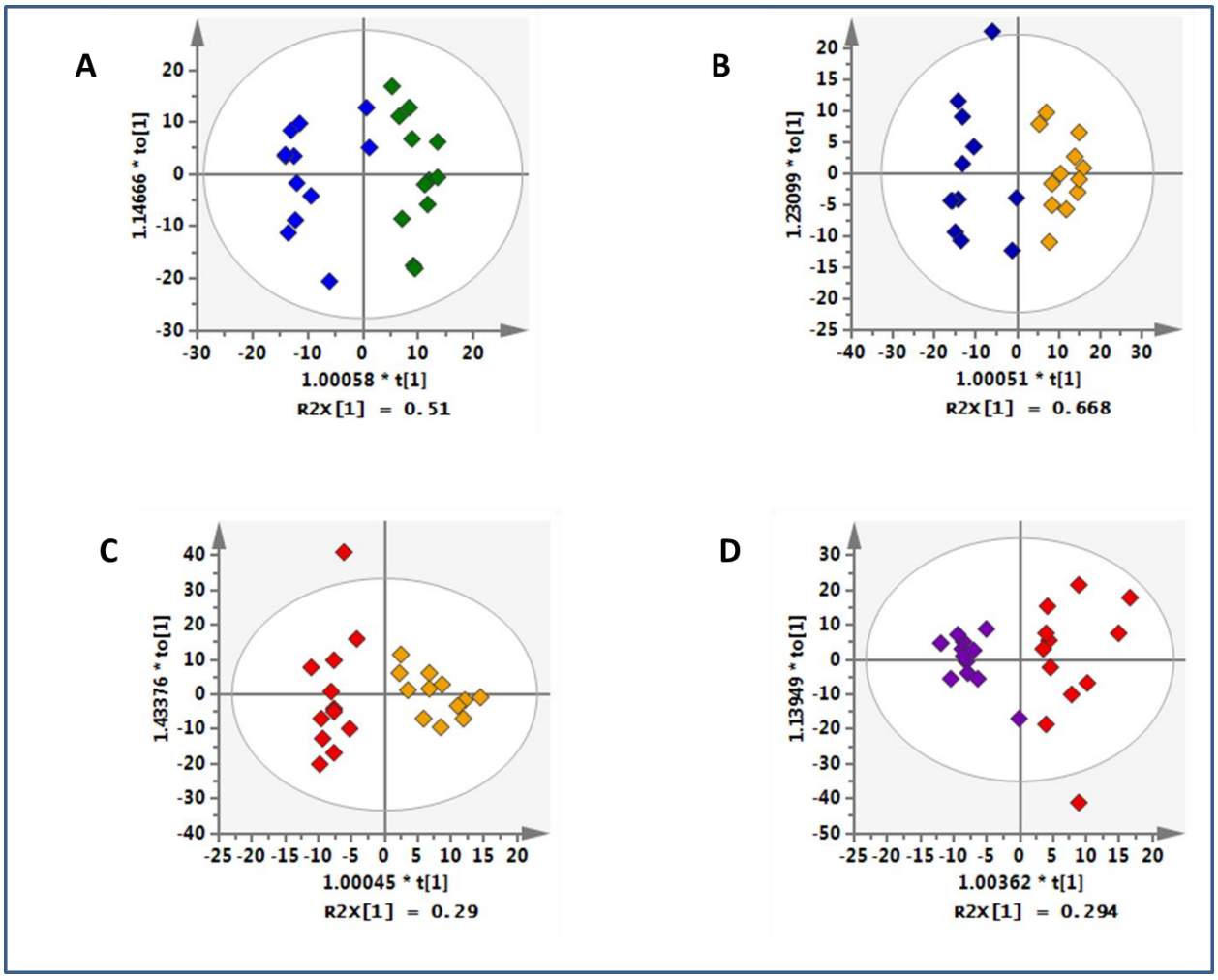
3.10 Supplementary material

Supplementary Table 3.1: Sampling dates in the three consecutive years

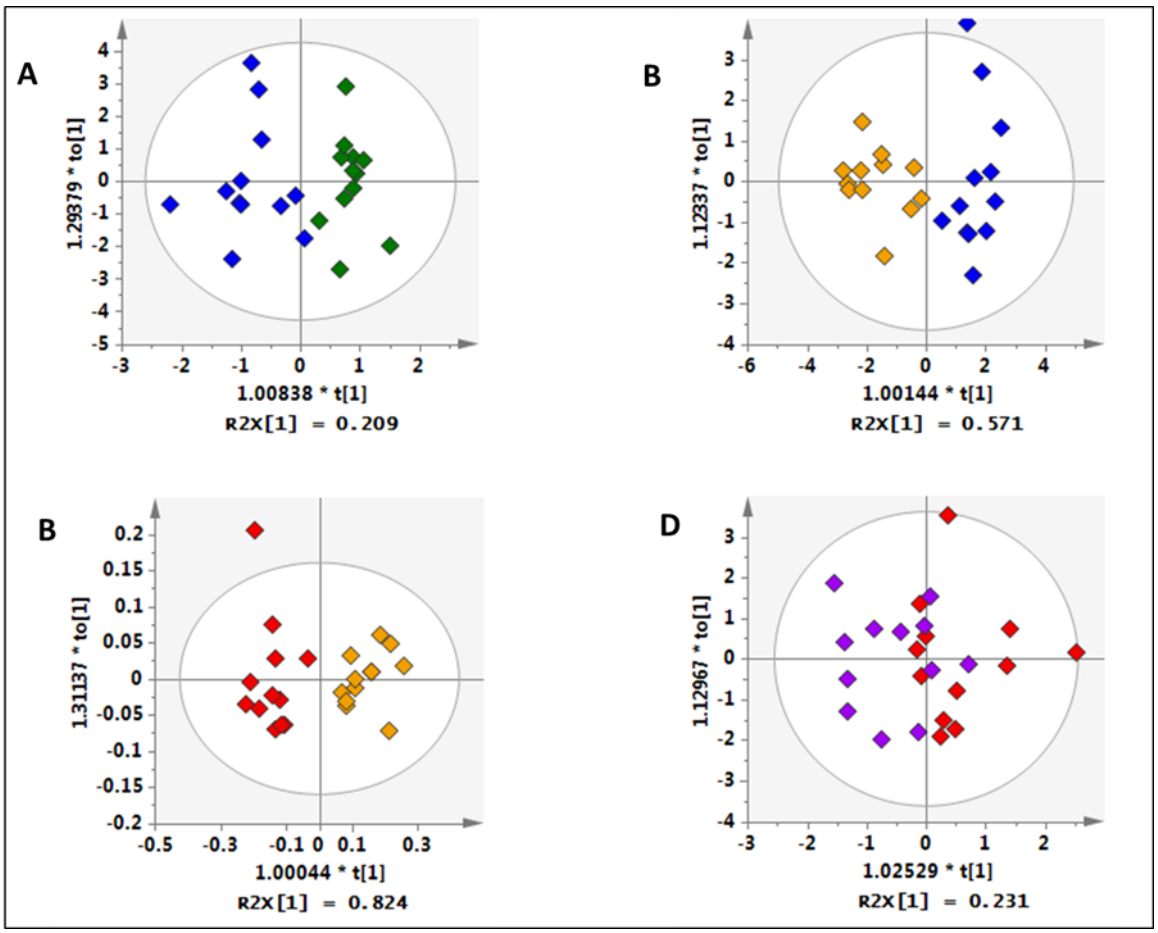
Developmental stage	Eichhorn-Lorenz phenological stage	2010/2011	2011/2012	2012/2013
Green	EL-31	19/12/2010	19/12/2011	21/12/2012
Pre-véraison	EL-33	12/01/2011	18/01/2012	18/01/2013
Véraison	EL-34	25/01/2011	05/02/2012	05/02/2013
Post-véraison	EL-35	22/02/2011	01/03/2012	01/03/2013
Ripe	EL-38	03/03/2011	13/03/2012	–



Supplementary Figure 3.1A: OPLS-DA scores plots for reference values as analysed at green/pre-véraison (A), pre-véraison/véraison (B), véraison/ post-véraison (C), and post-véraison/ripe (D). The colours represent the developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).



Supplementary Figure 3.1B:. OPLS-DA scores plots NIR spectra as analysed at green/pre-véraison (A), pre-véraison/véraison (B), véraison/ post-véraison (C), and post-véraison/ripe (D). The colours represent the developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).



Supplementary Figure 3.1C: OPLS-DA score plots and MIR spectra as analysed at green/pre-véraison (A), pre-véraison/véraison (B), véraison/post-véraison (C), and post-véraison/ripe (D). The colours represent the developmental stages: green (◆), pre-véraison (◆), véraison (◆), post-véraison (◆) and ripe (◆).

Supplementary Table 3.2: Average metabolite concentrations (mg/g FW) as determined by HPLC (i.e. reference data from frozen grapes), MIR (from frozen grapes) and MIR (from fresh berries).

Analyte	Sample	Green	Pre-véraison	Véraison	Post-véraison	Ripe
Tartaric acid	Reference	13.5	12	10.1	6.3	6.6
	Frozen	13.7	12.3	9.7	6.8	6.1
	Fresh	13.8	12.9	9.9	7.1	-
Malic acid	Reference	7.7	14.1	11.6	2.9	1.9
	Frozen	8.1	13.4	12.2	3.3	2.4
	Fresh	9.1	12.8	9.3	2.9	-
Succinic acid	Reference	4.3	2.9	4	2.9	1.3
	Frozen	4.4	3.3	3.6	2.5	1.6
	Fresh	4.6	3.5	4.4	2.5	-
Glucose	Reference	2.9	5.9	37.1	68.4	73.7
	Frozen	2.6	3.7	34.9	67.9	75.5
	Fresh	2.2	3.3	38.6	65.3	-
Fructose	Reference	1.3	2.3	31.3	67.3	73.5
	Frozen	1.2	2.4	31.9	65.8	71.7
	Fresh	1.3	1.8	34.8	67.1	-

CHAPTER 4

General discussion and conclusions, future perspectives and evaluation of the study

4.1 General discussion and conclusions

The unravelling of the grape berry development process is a daunting task that requires the application of different analytical approaches. Newer technologies like IR spectroscopy offer potential prospects for faster and efficient prospects for evaluating the quality of such an important crop as the grape berry. In order to face the many complex questions regarding grape berry development, powerful data acquisition and assessment techniques are needed. In this project IR spectroscopy in conjunction with chemometrics were used to evaluate grape berry development using NIR and MIR spectroscopy throughout the entire growth cycle. The combination of IR spectroscopy and multivariate data analysis is an upcoming powerful tool in the study of biological problems like those encountered in grape berry development.

The multivariate method used here provided simplified visual models that made data more transparent and thus facilitated data interpretation. In this thesis some useful data visualisation tools were presented and discussed to aid the interpretation of grape berry development in both qualitative and quantitative terms.

Time dependent grape berry models were described with PCA using MIR spectra. The stages that were modelled cover a wide range of key points of berry development including the period immediately after fruit set (EL 29-31), the period prior to ripening (pre-véraison) (EL 32-33), onset of ripening (véraison) (El 35), the period after the onset of ripening (post-véraison) (36-37) and fruit maturity (ripeness) (EL 38) (Coombe, 1995). The PCA approach showed the difficulties that can be encountered by using this technique on NIR spectra since the PCA models could not be used to visualise raw or pre-processed spectral data.

Application of the supervised OPLS-DA approach resulted in successful discrimination of grape berries according to developmental stage with both NIR and MIR spectra. It should be noted that the use of OPLS-DA could not discriminate between post-véraison and ripe stages with MIR data even though the NIR and MIR spectra were acquired from the same sample.

An extension of OPLS-DA, the S-line plot was applied to probe the key variables responsible for the observed differentiation of grapes according to developmental stages in the MIR region. The C-O stretch mainly in sugars was found to play a major contributory role in as far as stage discrimination was concerned. Still the observed correlation between C-O bond and stage discrimination does not suggest that sugars were solely responsible for stage discrimination.

In this study it was shown that using spectral data alone, PCA can be used to characterise grape berry samples according to developmental stage whilst supervised chemometric techniques such as OPLS-DA can be applied in the event PCA is incapable of doing so. Hence both NIR and MIR spectral data could be used in the prediction of grape berry developmental stage throughout fruit development.

A significant aim of the study was to investigate the possibility of using IR spectra to predict some metabolite concentrations at different stages of grape development during the berry growth cycle. It was observed that MIR models were superior to the NIR models in quantitative analysis. This observation on NIR would need additional probing since some sample presentation setbacks were experienced during NIR analysis *e.g* the frozen samples were not always homogeneous and in some cases the available sample could not completely cover the presentation surface (sample cup for rotating integrating sphere) resulting in increased light scattering.

Despite some observed differences between MIR and NIR spectroscopy in terms of qualitative and quantitative results, the two techniques were found to complement each other and for purposes like this project, they can be used in tandem since each one offers some unique strengths of its own. Looking at the literature, it is evident that FT-NIR has a predominant impact in the food industry in general due to its non-destructive nature and higher penetration into sample matrices. However ATR FT-MIR possesses enhanced spectral absorbance features that give it superiority in quantitative analysis (Gishen et al., 2005).

Our findings illustrated the usefulness of the PLS calibration models that were constructed from HPLC and MIR spectral data. The fitness for purpose of the respective models was very good as shown by the high coefficients of determination for glucose, fructose, tartaric acid and malic acid. The robustness of MIR models was tested by using spectra of samples whose reference data were known but had not been previously presented to the model. The low deviation (< 10%) between the reference and predicted results showed that ATR FT-MIR spectroscopic data and multivariate analysis provided a reliable and fast technique to predict targeted grape berry metabolite during berry growth. The predicted concentrations of sugars and organic acids were comparable to those presented by the reference method showing that the diamond ATR FT-MIR spectroscopic method can be relied on as an accurate prediction technique. The results showed that even compounds that are low in concentration in a grape berry sample like succinic acid can be accurately predicted by this technique.

A key feature of this NIR/MIR approach was that the study was used to design and execute a series of experiments that were capable of extracting multitudes of data using a few experimental runs. This thesis dealt with the qualitative and quantitative analysis of grape berries. The outcome from the study offered a scenario with a very good balance in cases where both the discriminatory power and the interpretation of key chemical changes related to biological class differences in grape berries are required. The described procedure offered good possibilities for the direct determination of glucose, fructose, malic acid and tartaric acid by MIR spectroscopy in complex matrices such as grape berries and can be considered as quite appropriate for the rapid analysis of these compounds simultaneously, providing a practical alternative to chromatographic methods when quicker results are needed.

Quantification of grape berry metabolites by IR spectroscopy is still a recently developed and growing technique for evaluating grape berry quality. IR calibration methods are curtailed by the fact that they cannot predict below the detection limits of the reference method. This hampers the technique as it can only be operational within the confines of the reference method.

The application of IR spectroscopy is already showing tangible evidence that grape berry analysis can be taken from the lab to the vineyard as shown by the development of portable hand-held instruments that are already available for on-plant evaluation of grapes (Barnaba et al, 2013; Giovenzana et al, 2014). Commercial application of IR spectroscopy in fruit grading lines was implemented at Japanese pack-houses, for sorting fruits since the mid-1990s (Nicolaï et al., 2007). That IR spectroscopy and chemometrics offer faster modes of grape berry analysis as required in the fruit and grape berry sectors is now a fact. The technique continues to improve as more sensitive IR spectrometers are being developed and some of them are equipped with highly sensitivity detectors (Nicolaï et al., 2007). This development has the potential to see the technique applied to metabolites occurring at very low levels in the berry such as flavour compounds. The huge advantage coming along with non-destructive approaches like IR spectroscopy is that in the fruit industry it allows for easier sorting, evaluation and general monitoring of the fruit products (Butz et al. 2005). The technique is likely to find significant demand from fruit producers, transporters and marketers.

As grape berry metabolites are in a constant state of flux, questions would remain as to the most effective benchmark for assessing them in the grape berry life cycle. IR spectroscopy has the speed and flexibility that offers a wide range of possibilities to answer some of these questions. Further research extending the use of IR

spectroscopy is likely to provide a significant shift in the way these issues are considered.

4.2 Future perspectives

IR spectroscopy and multivariate data analysis in grapevine research at present are still mainly focussing on the determination of the more abundant compounds like hexose sugars and a few organic acids like malic, tartaric, succinic and gluconic acids. There appears to be less concerted efforts to characterise the pentose sugars like arabinose and xylose which have a significant presence in grapes (Arnous & Meyer, 2009). The present research efforts on grape berry research are also dwelling more on primary metabolites. Apart from phenolics and related compounds secondary metabolites have not been so actively investigated by the IR spectroscopy/chemometrics approach.

IR spectroscopy has been known to be less suitable at analysing compounds occurring at levels below 0.2 mg/g (Bauer et al., 2008). In grapes most secondary metabolites including flavour compounds occur at levels much lower than the mg/g level. However this may be overcome as more sensitive IR spectrometers are being developed and some of them are equipped with highly sensitivity detectors (Nicolaï et al., 2007)

In regard to instrumentation there are remarkable advances with the use of mobile hand-held spectrometers especially with NIR applications (Barnaba et al., 2013; Giovencana et al, 2014). At the moment NIR spectroscopy remains at the forefront of non-destructive methods and appears to be at more advanced stages than MIR spectroscopy in regard to instrumentation and field application though the latter still has the edge in terms of better precision (Gishen et al, 2005). Such mobile instruments offer non-destructive techniques for on-plant analysis. With time it is most likely that mobile hand-held ATR FT-MIR spectrometers will also be developed.

The association of IR spectroscopy and chemometrics has established its own niche as an analytical tool and will continue to help researchers in evaluating more aspects of grapevine research. This technological combination is likely to give grapevine research impetus.

4.3 Evaluation of the study

With the benefit of hindsight, the following aspects could be considered unresolved aspects that could also form part of future work:

A better sample optimisation with FT-NIR would be necessary to give a bigger representative sample for NIR analysis than was carried out here. This could have offered a more informed performance of NIR spectroscopy in this study. PCA scores plots showed less separation between the sample clusters with frozen berry samples as compared to those generated from fresh samples. Another look at sample clustering at the five developmental stages from fresh grape spectra would need to be taken to assess if the co-clustering previously observed with frozen samples persisted or not.

PCA analysis of grape berry spectra was previously reported to show that grape berries clustered according to vintage (Barnaba et al., 2013). In our studies possible discrimination by vintage was noticed but could not be confirmed since we employed two different ATR FT-MIR spectrometers for different vintages. However, the PLS models derived from the same spectral data did not change with the machine change.

The results from this study showed immense potential in the application of IR spectroscopy for use in the classification of grapes by developmental stage and prediction of some individual sugars and organic acids but certainly further research including more grape varieties, different grape berry sites and increased number of cultivars are needed to test if the results obtained in this work can be extensively used for grape berry discrimination and berry metabolite prediction.

The thesis discusses methods and strategies to measure and evaluate grape berry in qualitatively and quantitatively. The author did not aim to develop the perfect qualitative and quantitative models for Sauvignon blanc grapes or let alone other grape varieties. However, it was hoped that the application of some of the strategies developed discussed in this thesis will help grape growers, oenologists and researchers in grapevine research.

4.4 Cited literature

Arnous, M., & Meyer, A.S. (2009). Quantitative prediction of cell wall polysaccharide composition in grape (*Vitis vinifera L.)* and apple (*Malus domestica*) skins from acid hydrolysis nonosaccharide profiles. *Journal of Agricultural and Food Chemistry, 57,* 3611-3619.

Bauer, R., Nieuwoudt, H., Bauer, F.F., Kossmann, K., Koch, K.R.,& Esbensen, K.H. (2008). FTIR spectroscopy for grape and wine analysis. *Analytical Chemistry, 1,* 1371-1378.

Barnaba, F.E., Bellincontro, A., & Mencarelli, F., 2013. Portable NIR-AOTF spectroscopy combined with winery FTIR spectroscopy for an easy, rapid, in-field monitoring of Sangiovese grape quality. *Journal of Agricultural and Food Chemistry, 94,* 1071-1077.

Butz, P., Hofmann, C., & Tauscher, B., (2005). Recent developments in noninvasive techniques for fresh fruit and vegetable internal quality analysis. J Food Sci. 70, 131-141

Coombe, B.G. (1995). Adoption of a system for identifying grapevine growth stages. *Australian Journal of Grape and Wine Research 1,* 100-110.

Giovenzana, V., Beghi, R., Malegori, C., & Civelli. R. (2014). Wavelength selection with a view to a simplified handheld optical system to estimate grape ripeness. *American Journal of Enology and Viticulture, 65,* 117-123.

Gishen, M.; Dambergs, R.G., & Cozzolino, D. (2005). Enhancing spectroscopy with chemometrics. *Australian Journal of Grape and Wine Research 11,* 296-305.

Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest Biology and Technology, 46,* 99-118.