Hinc patriam sustinet
Instituto Superior de Agronomia
Universidade Técnica de Lisboa

Vinifera EuroMaster

QUANTIFICATION OF CONDENSED TANNINS IN RED WINES BY FOURIER TRANSFORM MID-INFRARED SPECTROSCOPY (FTIR)

ALESSANDRO CORAZZOLA

Dissertação para obtenção do Grau de Mestre em

Vinifera EuroMaster – European Master of Sciences of Viticulture and Oenology

Orientador: Investigador Olga Laureano

Júri:
Presidente: Jorge Ricardo da Silva, Professor UTL/ISA
Vogais: - Emilio Cellotti, Professor, Universidade de Udine
- Olga Laureano, Investigador Coordenador, UTL/ISA
- José Carlos Rodrigues, Investigador, IICT

Lisboa, 2012

Acknowledgements

I acknowledge Professors Olga Laureano, Jorge M.Ricardo da Silva, José Carlos Rodrigues and all the technical staff from ISA, for their important help in this thesis.

Abstract

Grape-derived condensed tannins, also known as proanthocyanidins, are critical quality components for red wines. They have been subject of numerous studies in order to find a fast and reliable methodology for their quantification. In this work it has been tested the method using Fourier transform mid-infrared (FTIR) and chemometrics to quantify the amount of condensed tannins present in 88 different red wines, with the reference method of precipitation with methylcellulose. As well it has been provided a single laboratory validation of the method of fractionation of condensed tannins by reverse phase and quantification by reaction with vanillin. The models developed for the FTIR spectroscopy were not enough robust for the estimation of total condensed tannins, with low values of coefficient of determination and low RPD values (R^2 cross-validation: 0,76 and RPD cross-validation: 1,86). Validation of the fractionation method showed good performance in precision, with values of coefficient of variance for the three fractions FIII, FII and FI respectively of 5,2%; 11,4% and 11,6% and values of reproducibility of 168,1; 32,9 and 3,4 mg/L of epicatechin equivalents, but it was not possible to perform effective recovery studies.

Keywords: FTIR, condensed tannins, red wine, validation, reverse phase fractionation.

THESIS INDEX

Acknowledgements………………………………………………………………………………ii

Abstract……………………………………………………………………………………………iii

List of tables and figures……………………………………………………………………..vi

1. Introduction………………………………………………………………………………..1

1.1 Tannins: role and importance in enology……………………………………………………1
1.2 Condensed tannins…………………………………………………………………………….2
1.2.1 Extraction during wine production……………………………………………….3
1.2.2 Reactions in wine……………………………………………………………………….4
1.2.3 Potential effects on health of procyanidins……………………………………..5
1.2.4 Astringency and sensorial aspects related to condensed tannins………………………………………………………………………………..6
1.3 Methodology of detection…………………………………………………..…………7
1.3.1 Colorimetric methods………………………………………………….………….7
1.3.2 Gravimetric methods……………………………………………………………10
1.3.3 Chromatographic analysis……………………………………………………..…..11
1.3.4 Polymeric precipitation of tannins…………………………………………………12
1.4 Fractionation and isolation…………………………………………………………………..14
1.5 Infrared spectroscopy…………………………………………………………………..…..16
1.5.1 FTIR spectroscopy…………………………………………………………………18
1.5.2 Application of FTIR spectroscopy in enology……………………………….....20
1.6 Chemometrics…………………………………………………………………………..….24
1.6.1 Multivariate analysis……………………………………………….……………..25
1.6.2 Data Pre-treatment…………………………………………………………………25
1.6.3 Principal chemometrics techniques………………………………………………28
1.6.4 Explorative analysis………………………………………………..…………28
1.6.5 Multivariate regression……………………………………………..………29

2. Aim of the thesis……………………………………………………………………………..31

3. Materials and methods…………………………………………………………………..….32

3.1 Sample characterization………………………………………………………………..32

3.2 Chemical methods…………………………………………………………………………..33

3.2.1 Quantification of total condensed tannins by methylcellulose precipitation………33

3.2.2 Fractionation of proanthocyanidins by solid phase extraction…………………….34

3.2.3 Quantification of FI, FII and FIII fractions by vanillin reaction……………….…….36

3.2.4 Validation of analytical method…………………………..……………………….36

3.3 Fourier Transform Infra-Red methodology………………………………….………….38

3.3.1 Spectral acquisition of the samples……………………………………………….38

3.3.2 Development of calibration models………………………………………………….38

4. Results and discussion……………………………………………………………………..42

4.1 Quantification of total condensed tannins by Fourier Transform Infrared Spectroscopy……………………………………………………………………………….42

4.1.2 Models for prediction of total condensed tannins……………………………….………43

4.2 Single laboratory validation of the fractionation method developed by Sun et al…..……49

5. Conclusion……………………………………………………………………………………..51

6. References……………………………………………………………………………………..52

Annexes……………………………………………………………………………………..……60

List of tables and figures

Table 1: Applications of FTIR spectrometry in the region of MIR in the viticulture-enology…21

Table 2: Statistics of the prediction model without and with the two extreme values………..43

Table 3: Descriptive statistics of the red wine samples used for calibration and validation…45

Table 4: statistics of the model for the prediction of condensed tannins using FT-MIR and Methylcellulose precipitation assay……………………………………………………………..47

Table 5: Concentration of the three fraction of tannins for the reproducibility (mg/L Epic.E)..49

Table 6: statistical analysis of the reproducibility test…………………………………………49

Figure 1: Different types of molecular vibrations………………………………………………17

Figure 2: Schematic diagram of a Michelson interferometer, configured for FTIR…..………19

Figure 3: Data processing in FTIR……………………………..……………………………….19

Figure 4: Effect of normalization on near-IR spectra of five synthetic gluten and starch mixtures. Original spectra and spectra after normalization……………………………………..27

Figure 5: Fractionation procedure of proanthocyanidins by solid phase extraction……..…35

Figure 6-7: Red wines spectra in FTIR (3100-900 cm^{-1}) in overlay (6) and superimposed mode (7)……………………………………………………………………………………………42

Figure 8: PCA of the samples dataset………………………………...……………………..44

Figure 9: concentration of total condensed tannins estimated by FTIR spectroscopy instrument versus reference values from Methylcellulose method…………………………….45

Figure 10:concentration of total condensed tannins values predicted by the calibration model on the dataset versus the values from the method of reference on the validation dataset….46

1. INTRODUCTION

1.1 Tannins: role and importance in enology

In the last decades the world of enology has shown a growing interest in tannins, in order to increase the comprehension of their chemistry, analysis and biosynthesis. This interest is due by the fact that tannin content and composition are critical quality components of wines (Fernandez and Agosin, 2007). In the wine literature, the term „tannin" is used to identify a class of compounds that are formed in the grape berry. These tannins are extracted from the berry during winemaking and may be chemically modified as part of the winemaking/wine development process.

By definition tannins are substances able to originate stable combinations with proteins and other vegetal polymers (Handbook of Enology, 2006). In chemical terms, tannins are relatively bulky phenol molecules, produced by the polymerization of elementary molecules with phenolic functions. Their configuration affects their reactivity. They play an important role in different aspects related with the wine final quality: long-term color stability, haze formation and interaction with proteins, oxidation and browning, aging behavior, sensorial aspects like astringency, bitterness and color, anti oxidant and healthy properties.

Tannins can be divided in two class of molecules: hydrolysable tannins and condensed tannins. Condensed tannins are abundant in the different parts of the grape (seeds, skins and stems) and are extracted to wine during winemaking. Hydrolysable tannins are absent in grapes, and it is generally accepted that they pass to wine from oak during aging in barrels or added as oenotannins during winemaking (Clifford and Scalbert, 2000).

1.2 Condensed tannins

All grape-derived tannins having enological importance are condensed tannins (Herderich and Smith, 2005). They are referred also as proanthocyanidins, this term is due from the acid catalyzed oxidation reaction that produces red anthocyanidins upon heating PAs in acidic alcohol solutions. From the structural point of view condensed tannins are polymers formed through linkages, mainly C-C bindings, between flavan-3-ol molecules. In the specific: the most common covalent linkage occurs between the C4 of the pyran ring of an upper flavonoid with C8 of the A ring of a lower flavonoid. Also can occur bonding of flavonoids between C4 and C6 sites permitting branching of procyanidin polymer. Less common is the association combining C4 and C8 bonds as well as C2 and C5 or C7. Polymerization may also occur at other sites, depending on whether bonding occurs in association with oxidation, with acetaldehyde or glyoxylic acid. (Jackson, Wine Science 2008). The most important structural flavan-3-ol subunits forming these macromolecules are (+)-catechin, (-)-epicatechin, but several other flavan-3-ol compounds participate as building units of condensed tannins like: (-)-epigallocatechin, epicatechin-3-O-gallate, (+)-gallocatechin, (+)-gallocatechin-3-O-gallate and (-)-epigallocatechin-3-O-gallate. Basic catechin units may not be considered as tannins, as their molecular weight is too low and they have very restricted properties in relation to proteins. They have only enough high weight in the dimeric form (Handbook of enology). The size of the polymer is often described by the degree of polymerization (DP), that represents the number of flavan-3-ol molecules in the polymer. So following the increasing polymerization degree, we can find dimeric, trimeric, oligomeric and polymeric procyanidins. Frequently, fractionation steps produce fractions that contain polymers of several different lengths. In this case, the average or mean degree of polymerization (mDP) is used.

Condensed tannins are present in all the solid part of the grapes: skins, seeds and stems, but with structural differences among these parts. In some cases they have been detected also in the pulp, mainly in “teinturier” varieties. There is also a considerable variation in the types and concentration among cultivars (Kovac *et al.,* 1990), and is well established that there is a variation due to the vintage, cultural practice and environment around the developing fruit (Kennedy J.A. 2008).

Grape tannins are synthesized during berry formation and are the results of the secondary metabolism of the vine plant. This reaction is essentially completed by veraison prior to the commencement of anthocyanin biosynthesis. Hence, the largest amounts of tannins per berry can be extracted at veraison, followed by a reduction in extractable tannin as grape

maturity progresses (Kennedy *et al.* 2000, Downey *et al.* 2004).This period is called "tannin maturation" in which there is an apparent decline in condensed tannin concentration. While this is in part due the increasing volume of the berry, with consequent dilution of tannins synthesized pre-véraison, there is also a marked decline in the total amount of tannin that can be extracted chemically from skins and seeds (Scollary, 2010).

The biosynthesis of the phenolic compounds that are important to wine quality share a common pathway. They are, in one way or another, formed via the shikimic acid pathway, also known as the phenylpropanoid pathway, but proanthocyanidin production in plants is still incompletely understood. While the genes and enzymes responsible for the biosynthesis of the monomeric subunits catechin and epicatechin have been identified (Tanner *et al.* 2003), details of the polymerization reactions that yield condensed tannins remain unknown. Clearly, the unresolved biosynthesis of tannins resembles one of the last frontiers in plant secondary metabolism (Herderich and Smith, 2005).

1.2.1 Extraction during wine production

Once grapes are harvested phenolic composition in the wine becomes dependent upon processing in the winery. So the condensed tannins composition in the wine depends not only on the grape composition, but also on their extraction and consequent reactions after wine-making and ageing process. Condensed tannins present in the skin and seeds make up the majority part, with stem derived making a minor component if included. Generally speaking there are different factors that affect maceration and so the extraction: time, temperature, sulphur dioxide, ethanol, enzymes, cup punching programs, solid/liquid proportion. In this way the wine composition is influenced by the duration of pre- and post-fermentation maceration phases and by treatments enhancing cell wall or berry degradation (Wine chemistry, 2009).
White wines content of these compounds are quiet low, but an increase can be obtained by procedures like skin contact maceration, sulphur dioxide addition and clarification degree. For red varieties the attention is focused not only to the total rate of extraction but also to the seed or skin origin of the condensed tannins. In the final wine, the skin and seed tannin proportions are generally different than those found in the berry, considering the different composition of the two types of tannins modifying the proportion of them in the wine means also modifying the final quality of the wine. Skin tannins are generally extracted early in fermentation and as the maceration time increases, the rate at which seed tannin are extracted increases (Peyrot des Gachons and Kennedy, 2003). Tannin extraction will

increase throughout fermentation (Ribéreau-Gayon *et al.*, 1970) and therefore, at some point seed tannin dominates the tannins present in the wine.

1.2.2 Reactions in wine

Reactions in condensed tannins composition start as soon as the beginning of wine making and continue in fermentation and ageing process. These changes involve both enzymatic and chemical reactions. The first processes, due to grape, yeast or fungi and exogenous enzymes occur mostly in the early stages, while chemical reactions continue during ageing. (Wine science, 2008).

The procyanidin molecules from the grapes tend to polymerize, condense with anthocyanins and combine with plant polymers such as proteins and polysaccharides (Handbook of Enology, 2006).

The structure of procyanidin presents in an acid environment a great reactivity caused by the acid hydrolysis of a C-C bindings with the formation of a procyanidin with a low molecular weight and a carbocation with electrophilic character. This molecule presents a huge reactivity and can bind different nucleophilic compounds, like another procyanidin, and increase in this way its degree of polymerization.

Another characteristic of the procyanidins is the oxidizability. Light and temperature, as well as the presence of hydroperoxides and certain metals, promote the formation of oxygenated radicals (Waters, 1964). Condensed tannins react more-or-less easily with free radicals, according to their configuration. These reactions lead to polymers and insoluble brown pigments.

In the development of the colour during the wine aging, condensed tannins reactions perform an important role. The reactions of tannins with anthocyanins are responsible for the formation of pigmented polymers and so they perform an intensification and stabilization of the red wine color. These condensations can come off different mechanisms and may produce compounds with various characteristics, according to the type of bonds.

A wide range of mechanisms have been proposed to explain the formation of stable pigments. These include anthocyanin-flavanol (A-F) direct condensation reactions, flavanol-anthocyanin (F-A) direct condensation reactions, acetaldehyde mediated anthocyanin-flavanol and anthocyanin-anthocyanin condensation reactions, pyranoanthocyanin formation and flavanol-flavanol condensation reactions. A more detailed understanding of the mechanisms of the reactions that are associated with pigment evolution is required (Scollary, 2010).

Three main type of reactions regarding condensed tannins can be distinguished:

- Direct condensation reactions Anthocyanins-tannins, in which the anthocyanins in cationic form react with procyanidins;
- Condensation of tannins with anthocyanins in which the carbocation formed by procyanidin reacts with the nucleophilic sites of anthocyanins molecule as carbinol basis;
- -Indirect reaction: condensation with an ethyl cross-bond. Ethanal forms a carbocation that reacts with the negative nodes of the flavanols (procyanidins, catechins), and then with the anthocyanins in carbinol base form (Handbook of Enology, 2006).

Tannins have the property to associate and interact with proteins and polysaccharides, in order to have stable combinations. Especially the interactions between tannins and proteins have been extensively studied, owing to their role in haze formation, astringency perception and nutritional and anti-nutritional effects resulting from inhibition of various enzymes and reduction of dietary protein digestion (Wine science, 2008). Polysaccharides have the capacity to affect protein-tannins interaction and the astringency response in wine (Ozawa *et al.*, 1987; Luck *et al.*, 1994).

1.2.3 Potential effects on health of procyanidins

At this time, the study of the proprieties of the procyanidins presents a great interest due to the pharmacological and nutritional implications. In fact the moderate consumption of wine, especially red, is associated with positive effects on human health. Numerous studies has been done on this field but we are still far from a complete understanding on the mechanisms and molecules closely involved in the phenomenon. In particular procyanidins may be one of the polyphenolic component that ensure positive benefits. Procyanidins have been identified as high performance, low toxicity and highly bio-available. A steady stream of animal and in vitro studies supplemented by epidemiological evidence and human studies reveal numerous health benefits associated with these compounds.

Chief among the benefits is antioxidant protection against heart disease and some cancers. Oxidative stress is involved in many diseases, such as atherosclerosis, Parkinson's disease, heart failure, myocardial infarction, Alzheimer's disease and chronic fatigue syndrome. Procyanidins have been found as really powerful antioxidants, they slow or prevent the oxidation of other molecules by capturing free radicals. These properties revealed itself even more powerful than other traditional molecules able to capture radicals (Da Silva R., 1995).

Oxidized low-density lipoprotein cholesterol damages the cells that line blood vessel walls by provoking numerous responses including inflammation, smooth muscle cell proliferation and

clotting mechanisms, all of which lead to atherosclerosis. Procyanidins develop its potential benefic effect at the level of the main factors involved in this illness: protection from the degradation of structural proteins (collagem above all), depuration of cholesterol and protection from the excessive permeability of arteries.

Other studies demonstrated that procyanidin B dimers in red wine could be used as chemo preventive agents against breast cancer by inhibiting the conversion of androgens to estrogens in breast tissues (Shiuan Chen, Ph.D.). Moreover procyanidin can induce apoptosis (self-destruction of cancer cells) and necrosis (death of cells as a result of an outside agent) of prostate cancer PC-3 cells in a mitochondrion-dependent manner (X J. Shang *et al.*, 2008) and has been reported that procyanidin has multiple anticancer effects on different cancer cells (it is identified as being toxic to cancer cells), such as cutaneous carcinoma, oral carcinoma, breast carcinoma, bronchogenic carcinoma, liver carcinoma, prostate carcinoma, pancreatic carcinoma, gastric carcinoma, and more, along with growth promoting effects on normal cells.

Procyanidins may also protect against viruses like Herpes, venous insufficiency and protect the body from toxins.

1.2.4 Astringency and sensorial aspects related to condensed tannins

One of the most studied effects of condensed tannins is their impact on sensory perception of the wine, not only for astringency but also for the mouthfeel and other tasting notes. A considerable amount of literature can be found about this topic, here we are just going to pass through the most important steps.

First sensation linked with tannin quantity and composition is astringency. It is a tactile sensation, which is described in sensory terms as drying, roughing or puckering (Lee and Lawless, 1991). It is generally considered that we observe this as a loss of lubrification due to the tannins binding and precipitating our salivary proteins.

Although it is widely accepted that astringency results from the interaction between condensed tannins and salivary proteins, Santos-Buelga and de Freitas, (2009) note that the detailed physiochemical mechanism needs further refinement. In fact there are several matrix components (like ethanol, acidity, viscosity, simple sugars, polysaccharides, anthocyanins, etc.) in the wine medium that influences this protein-tannin mechanism; sometimes they can oppose to the reactivity of tannins decreasing their aggressiveness, sometimes increasing it.

In addition condensed tannin structure/function activity must be deeper understood especially regarding these three cardinal points:

- Correlation between Degree of Polymerization of condensed tannins and sensorial perception of astringency.
- Influence of tannin modification during wine ageing on sensory perception.
- Contributions of condensed tannins from skins and seeds, and degree of maturation to the perception of astringency.

Another important sensation linked with tannin presence is bitterness. It is a taste response linked to specific receptors on the tongue. The response is generally assumed to be restricted to small molecules; that is the molecules require the appropriate size and shape to interact with the receptor site (Cheynier *et al.*, 2006).

Tannins are supposed to be connected with other sensation like body, structure, volume, thinness and harshness but still a lot of work must be done in this way.

Regarding these points not only an analytical development is researched in order to have precise and correct method of analysis and fractionation, but also a large effort must be done in connecting these laboratory data with the real perception in the mouth of the wine taster. In fact with the ability to measure tannins and their sensorial impact comes the potential to modulate the tannin profile of red wine to a particular style or market.

1.3 Methodology of detection

There are a wide range of measurement techniques applied to the determination of condensed tannins in grape homogenates, grape skins and seeds and as well in wines. The measurement capacity ranges from the simple UV spectral measurement at a single wavelength to high cost sophisticated instrumentation such as nuclear magnetic resonance (nmr), LC-mass spectrometry (LC-MS) and LC-MS-MS (tandem mass spectrometers) (Scollary, 2010).

In summary, the methods can be broadly categorized into measurements based on colorimetric derivatisation, gravimetric analysis, chromatographic analysis and polymeric precipitation of tannins. In the next paragraphs we are going to resume the various possibilities and their specificities, following the recent reviews papers by De Beer *et al.* (2004), Herderich and Smith (2005), Cheyner (2006), Seddon and Downey (2008) and Scollary (2010).

1.3.1 Colorimetric methods

Colorimetric methods are abundant in the published literature. The basic mechanism is a reaction of tannins with a chemical to yield a colored complex, which is measured spectrophotometrically. This method works well with purified samples, and yields method

specific results, but analysis of wine is hampered by interferences with other phenolic compounds (Herderich and Smith, 2005). Indeed it is generally time consuming, use chemicals that require special handling and, most important, lack selectivity for tannins. In the end we can say that now available there are several wet chemical methods that provide more accurate and precise information than can be obtained through colorimetric methods (Scollary, 2010).

One of these classic colorimetric methods is the buthanol-HCl assay. In hot solution of butanol and hydrocloridric acid, it provides the depolimerization of the condensed tannins to yield single subunits, which are then oxidized to red coloured anthocyanidins. Then the absorbance is measured at 550nm wavelength (Porter *et al.* 1986). This method is selective for proanthocyanidins and does not detect hydrolysable tannins.

On the other hand it presents several drawbacks: it is influenced by variations in the water and metal content of the assay mixture, the transformation of PA into anthocyanidins is not complete (Scalbert, 1992) and side reactions are common during the transformation and lead to the formation of red-brown polymers absorbing around 450 nm (Scalbert, 1992). These factors together provoke errors of estimation, so it must be taken with caution as a quantitative assay.

Other methods, such as the Prussian Blue or the Folin-Ciocalteu assays, are based on oxidation-reduction reactions in which the polyphenol is oxidized. As above, also these reactions are non-specific, numerous polyphenols can react in addition to tannins and several interferences with wine compounds has been found. Although these two methods are widely used the results are not enough robust in order to be completely trusted.

It has been published, by Stevanato *et al.* in 2004, also an enzymatic method based on the peroxidase-catalysed oxidation of phenols to phenoxyl radicals, which can react with aromatic substrates to form colored adducts. The final quantification is done measuring the absorbance at 500 nm wavelength. In comparison with the widely used Folin-Ciocalteu method, this method appears to be more specific and more rapid and as a whole is not affected by the common interfering substances such as ascorbate, citrate, and sulfite (Stevanato *et al.*, 2004).

Thus, Vanillin assay and dimethylaminocinnamaldehyde (DMACA) assay are the two colorimetric methods mostly used for quantification of total proanthocyanidins because of their sensitivity, specificity and simplicity; though several factors can affect their precision and accuracy for which precaution should be paid (Sun *et al.*, 1998b). Mainly they rely on the formation of colored products from the reaction between tannins and the aldehyde reagent.

For these two types of assay the authors showed different position, some taking one side some the other side. The DMACA employs the absorbance at 635 nm of the adducts formed

between the aldehyde 4-dimethylammino-cinnamaldehyde and phenols such as tannins (Treutter, 1989). Some authors think it is more specific than vanillin but still monomeric polyphenols will react causing a risk of tannin content overestimation (McMurrough and McDowell, 1978). DMACA assay has been proposed to asses oxidative ageing potential of white wines (Schneider, 1995) and for post column derivatisation during HPLC analysis (Treutter *et al.*,1994; de Pascual-Teresa *et al.*, 1998).

Vanillin was used for many years as standard colorimetric method but lack in reproducibility has been often reported. The vanillin reaction involves reaction of an aromatic aldehyde, vanillin with the metasubstituted ring of flavanols to yield a red colored extract with maximum absorbance at 500 nm (Price, van Scoyoc and Butler, 1978). In order to improve it Sun *et al.* in 1998 studied some factors mostly affecting the vanillin assay, to ensure its precision and accuracy. Since this method has been used in this thesis we are going to speak more deeply about every factors interested.

a) Influence of acid nature and concentration

The vanillin reaction is carried out with acid medium. Considering hydrochloric acid and sulfuric acid, for the same acid normality, the *A*500 obtained by using H2SO4 was much higher than that obtained by using HCl. So H2SO4 showed a better sensitivity. Regarding the concentration moreover for total analysis of catechins and PA, we should use H2SO4 with a concentration of nearly 10%; in this case, (+)-catechin could be taken as reference standard, and the measured value should be approximately true concentration in sample.

b) Influence of reaction time

It is recommended to fix the reaction time at 15 min for catechin estimation. However, for proanthocyanidin estimation, the maximum *A*500 should be taken as the measured value, correspondent to the maximum value of the curve of absorbance during the reaction time.

c) Influence of temperature

Because increasing temperature increased the *A*500 value, 25-35 °C is recommended for catechin estimation, and it should be well controlled. Instead there is no problem for PA estimation when the temperature was not significantly changed, and this estimation could be performed at room temperature.

d) Influence of water content

Because the *A*500 also depends on the acid concentration or, more properly said, the pH, a change of water content modifies undoubtedly the *A*500. So absolute methanol and reaction model free of water seems the best solution for stabilizing the color and obtain the best sensitivity.

e) Influence of Vanillin concentration

The vanillin, as reagent, must be excessive so that the reaction is complete. Best ranges of vanillin concentration between 10 and 12 g/L are recommended, because higher concentrations can give the formation of colored self-aggregates.

f) Effect of diffused sunlight

Analysis of variance showed that, both for (+)-catechin and for oligomeric PA, there were no significant differences at the 5% level between vanillin reaction undergone in the dark and vanillin reaction undergone in diffuse sunlight. So experiments should be performed in diffuse sunlight.

g) Interference substances

Ascorbic acid or ascorbate interfered with the vanillin assay and should be preliminarily separated from catechins and Proanthocyanidins. Also Chlorophyll from grape stem extract can interfere and should be extracted by hexane .It has been also reported that in acidic conditions, anthocyanins absorb at 490-540 nm, which coincides with the colored product (500 nm) of the vanillin assay. This interference can be simply eliminated using a suitable blank that is the same as the reaction medium, but in the absence of vanillin (Broadhurst and Jones, 1978).

1.3.2 Gravimetric methods

Gravimetric methods are based on the capacity of metals, such as the trivalent cation of the rare earth metal ytterbium (Yb^{3+}), to complex and precipitate tannins (Reed *et al.,* 1985). The advantage of a gravimetric procedure is that standards are not needed. But still there are some problems linked to the fact that not all polyphenols are precipitated and low repeatability in wines with low levels of tannins. (http://www.ansci.cornell.edu/plants/toxicagents/tannin.html)

A gravimetric adaptation of precipitation with polyvinylpyrrolidone (PVP) has also been reported (Makkar *et al.*, 1993), in which the dried mass of tannin extract before and after precipitation is used for quantification (Sarneckis *et al.*, 2008). The method represents an improvement over the gravimetric one of Reed *et al.* 1985, which lack specificity for tannins This method is not very sensitive and tends to underestimate tannins. However this assay can only be applied to extractable tannins.

1.3.3 Chromatographic analysis

A range of methods using expensive, high maintenance equipment (best suited to use in research facilities) has been developed for analysis of condensed tannins. These are primarily HPLC (High-performance liquid chromatography) based methods, including normal phase (Rigaud *et al.*, 1993), reversed-phase (Cozzolino *et al.*, 2004) and size exclusion chromatography columns (Yanagida *et al.*, 2003) coupled with mass spectrometry, diode array or electrochemical detectors.

Normal phase HPLC uses a polar column material (stationary phase) so that non-polar compounds are eluted early followed by polar compounds. There has been some success with this technique, but is requires specific workup for each application to ensure the successful separation of target analytes (condensed tannins) from other UV absorbing compounds such as anthocyanins (Scollary, 2010). Resolution by normal-phase HPLC is much poorer than that by reverse-phase HPLC and overlapping becomes more important as molecular mass increases (Sun and Spranger, 2005).

Reverse phase chromatography uses a non-polar phase and has been highly successful in separating the products obtained after depolymerization in the presence of so-called nucleophilic agents. It permits to separate low molecular weight proanthocyanidins, because oligomeric and polymeric forms are co-eluted with a large peak without resolution (Rigaud *et al.*, 1993).

MS (Mass Spectrometry) and nmr (Nuclear magnetic resonance) are ideal „partner tools" for HPLC as they provide detail of the mass and arrangement of atoms (i.e.: the structure) in the molecule, often in combination with a DAD (Diode Array Detection) detector.

Mass spectrometry is an ideal method for determining the molecular mass of a compound that allows an empirical/molecular formula to be estimated. There are a range of ionization methods used in MS, all with advantages and limitations. For a compound to be detected by MS, it must be charged as the signal that is obtained is actually the mass/charge ratio. Electrospray ionization mass spectrometry (ESI-MS) has been proved to be a very powerful tool for characterization of proanthocyanidins, in particular for detection of individual oligomeric and polymeric proanthocyanidins in a mixture or in heterogeneous solutions, providing the molecular mass, number of constitutive moieties and substituents (Guyot *et al.,* 1997; Le Roux *et al.,* 1998; Hayasaka *et al.,* 2003).

Nuclear magnetic resonance (nmr) is a potentially ideal technique for assessing structural characteristics of condensed tannin extracts (Cheynier and Fulcrand, 2003). For nmr to be successful however, the molecule must be pure, both chemically and isomerically. This requirement often limits the application of the methodology. One evolving technique for

general compound analysis can be HPLC with an nmr detector (Wolfender *et al.*, 2003).The use of chemometrics for the analysis of instrumental data will probably also give some nmr developments.

Considering all these techniques, the major limitation is the absence of standard reference materials. This must be rectified as a matter of urgency (Scollary, 2010).

1.3.4 Polymeric precipitation of tannins

Tannin measures that are based on selective precipitation by polymers represent an interesting and viable option for routine applications which do not require access to HPLC systems. Indeed protein precipitation functionally defines tannins, so that by definition protein precipitation is the most appropriate analytical principle (Sarneckis *et al.*, 2006). Starting from the use of blood from a freshly pricked thumb (haemoglobin) proposed by Bate-Smith in 1973, numerous protein were tested. A gelatin index has also been proposed by Glories in 1984, whereby tannins are ranked by their propensity for precipitation of gelatin, but the variability in composition of gelatin gave problems of reproducibility. One of the most widely used for rapid condensed assay is the Adams-Harbertson assay (2003), an expansion of the protein-based precipitation method that uses bovine serum albumin (BSA), where the precipitation step was complemented with the bleaching effect of sulphite. Far from a unified judgment upon the robustness of this method, it has its supporters and detractors. In fact while Mercurio and Smith (2008) reported acceptable precision (measured as repeatability) for the Adams-Harbertson assay, Brooks *et al.* (2008) have published a damning report of the Adams-Harbertson method's precision. Brooks argued that significant underestimation of the tannin content of wine would occur if it is only used the Adams-Harbertson assay for monitoring the winemaking process.

In general, all protein precipitation assays are dependent on many variables, that are potential sources of analytical variation. Nevertheless, all protein precipitation assays are hindered by our inability to measure directly the removed tannin, as the ideal absorbance for spectral quantification of tannin, at 280 nm, suffers interference from the added protein precipitant (Herderich & Smith, 2005). Instead a positive aspect of the protein precipitation assays is that the interaction of proteins with tannins can be used to model astringency perception in humans (Bate-Smith, 1973).

In order to obtain a direct spectral measure of tannin, and to circumvent the non-specificity of most colorimetric derivatisation methods, we turned to precipitation by non-proteinaceous polymers. Three non-proteinaceous polymer precipitants that had been previously reported were considered as candidates, namely polyvinylpyrrolidone or polyvinylpolypyrrolidone (PVP or PVPP) (Makkar, 1993; Antoine *et al.*, 2004), polyethylene glycol (PEG) (Jones, 1965; Jones and Mangan, 1977; Makkar *et al.*, 1995; Silanikove *et al.*, 1996) and methyl cellulose (MCP) (Montedoro and Fantozzi, 1974) (Sarneckis *et al.*, 2006). PVP and PVPP showed problems of selectivity and detection due to the capacity to complex with polyphenols others than tannins. PEG is exploited for tannin analysis in plant forages, showing simplicity and accuracy (N. Silanikove *et al.*, 2001) but it has not been applied to grape or wine matrices.

MCP method which provides the precipitation using the polysaccharide methylcellulose, was developed at AWRI (Sarneckis *et al.*, 2006) and revised by Mercurio *et al.* in 2007. Precipitation of tannin by a polymer with minimal absorbance at 280 nm was selected as the analytical basis for the assay. This avoided the need for indirect quantification methods, such as radioactive tracers or non-specific colorimetric assays. Considering also the non-toxic, food-grade nature of methyl cellulose, it was selected as polymeric precipitant. The methyl cellulose precipitable (MCP) tannin assay is based on a measure of 280 nm absorbance of phenolics, before and after precipitation of tannin by methyl cellulose. This tannin assay enables the complete precipitation of tannin from red wine and from the grape homogenate extracts studied. It has proved to be selective for condensed tannins and did not suffer interference from other 280 nm-absorbing phenolics such as anthocyanins or catechins.
We have also to consider that the results might reflect the analytical method as much as they reveal sample composition, and it is difficult to compare tannin data that have been obtained by different methods. In fact there has not been a 'round-robin' comparative study of the MCP assay, as an interwinery comparative study. Sarneckis, Dambergs, Jones, Mercurio, Herderich & Smith in 2006, focused a study for the optimization and validation of this method. Since we used the MCP method in this thesis we are going to see through the main points of this study:

a) <u>Effect of pH, ethanol concentration and ammonium sulphate on assay performance</u>

pH value (3.2-4.5 and 7.4), ethanol (10%, 25% and 50%) and ammonium sulphate concentration increase, within the above ranges do not appear to affect the performance of the methyl cellulose assay.

b) Effect of methyl cellulose volume, concentration and composition

The general trend observed by increasing the methyl cellulose concentration from 0.04% to 0.4% (w/v) or by increasing the volume of methyl cellulose solution was an increase in the precipitation of tannin and subsequently an increase in A280 (tannin). Nevertheless it was found that 0.04% methyl cellulose solution removed no anthocyanins while 0.4% methyl cellulose solution removed up to 5% of the more hydrophobic anthocyanins (e.g. malvidin-3-glucoside-6'-coumarate). Thus 0.04% methyl cellulose solution was chosen.

c) Effect of standing time on assay results

The most consistent result were found when the assay was left at room temperature without centrifugation and after 10 minutes in this conditions.

d) Assay validation for red wine:

There was good correlation between tannin measured by reverse-phase HPLC and the MCP tannin assay for 121 Australian red wines (coefficient of correlation $r = 0.74$ and coefficient of variation CV= 3.0%).

1.4 Fractionation and isolation

For the characterization of tannins, they must be primarily extracted from samples of grape or wine, before they can be isolated and fractionated by extraction in solid phase mainly by reverse phase, but also by liquid-liquid extraction (Cheynier *et* Fulcrand, 2003). Separation and fractionation of condensed tannins and polyphenol-derived compounds into different sub-fractions are generally a first step for their further identification and characterization. In fact to conduct further research, it is necessary to have some methods that allow separation of tannins according to their size for analytical and preparative purposes (Cedric Saucier, 2001). Fractionation of red wine polyphenols have been always a challenging task for scientists due to the structural diversity and complexity of these molecules. One of the major problems actually encountered is the difficulty to separate oligomeric proanthocyanidins from anthocyanins and polymeric proanthocyanidins from pigmented polymers (Sun *et* Spranger, 2005). From the literature many methods have been proposed to separate PA according to their Polymerization Degree. Most of these methods are based on the separation by gel permeation chromatography (GPC), solidphase extraction (SPE) on C18 cartridges or more recent techniques such as countercurrent chromatography. Concerning the analytical aspect, convenient separation can be achieved by normal-phase TLC or HPLC. For semi - preparative, preparative or qualitative purposes, gel chromatography with different gels has been used. Unfortunately, only oligomers up to five are easily separated with these methods,

and irreversible absorption often occurs which limits the life of these expensive gels. A number of liquid chromatographic methodologies have been described in the literature for fractionation of polyphenols using Sephadex G-25, Sephadex LH-20, Sepharose CL-4B, Fractogel (Toyopearl TSK-HW 40(s), Fractogel (Toyopearl) TSK 50(f), and inert glass microparticles (Ryszard Amarowicz, 2006). Considering the aim of this introduction we are not going through all the details and method of these several techniques, but we will underline their impact and principal drawbacks on the analysis. The main shortcomings of these techniques are that they are very delicate and time consuming, which makes difficult their use for routine analysis (Sun, 1998). Liquid-liquid extraction with ether to isolate monomeric flavanols and with ethyl acetate to isolate oligomeric proanthocyanidins still is used but the recovery is poor and ethyl acetate extracts also a small amount of anthocyanins.

Solid-phase extraction (SPE) is a rapid and easy method which permits quantitative separation of polyphenols into various classes (Jaworski and Lee, 1987; Oszmianski *et al.*, 1988; Revilla *et al.*, 1991; Sun *et al.*, 1998a), with significant reduction of solvent used. It is an extraction technique based on the selective portioning of one or more components between two phases, one of which is a solid sorbent. The second phase typically is a liquid, but it may be also an emulsion, gas or supercritical fluid. Once equilibrium has been reached the two phases are separated by decanting, filtration, centrifugation or a similar process. If the desired analytes are absorbed on the solid phase, they can be selectively desorbed by washing with an appropriate solvent. If the components of interest remain in the liquid phase, they can be recovered via concentration, evaporation, chromatographic separation and recrystallization. The real benefits of SPE are gained when the sorbent is packed efficiently into a tube to form an uniform bed. By passing through this bed the solid-liquid extraction technique becomes a form of column chromatography (P.D. McDonald, 2001). This is a rapid, economical, and sensitive technique that uses several different types of cartridges and disks, with a variety of sorbents. Sample preparation and concentration can be achieved in a single step. Interfering sugars can be eluted with aqueous methanol on reversed-phase columns prior to elution of flavonoids with methanol (A. Marston *et* K. Hostettmann, 2006). This method was particularly analyzed by Sun in different studies. SPE on Sep-Pak C18 cartridges can used for quantitative fractionation of grape and wine proanthocyanidins according to their degree of polymerization. The wine or grape extracts could be separated on C18 Sep-Pak cartridges into three fractions by different organic solvents, namely: 3-flavonols monomeric (FI), oligomeric proanthocyanidins (FII) and polymeric proanthcyanidins (FIII). This cartridge permits the removal of sugars and others highly polar compounds like organic acids, amminoacids and proteins. The main limitation of SPE cartridge is the

incapacity to separate the anthocyanins polymeric proanthocyanidins from polymeric pigments (Cheynier *et* Fulcrand, 2003).

1.5 Infra-red spectroscopy

Infra-red spectroscopy is an analytical technique widely used for the identification and quantification of organic, and in some cases inorganic, materials.

The term spectroscopy refers to the study of the interaction between material substance and, in general terms, electromagnetic radiation. In fact electromagnetic radiation is a form of energy which propagates itself in an ondulatoric way through the space. The whole of electromagnetic radiation is called electromagnetic spectra. If the nature of the electromagnetic radiation is fundamentally always the same, different forms can be divided within the spectra according to the three correlated values of wavelength, frequency and energy; so we can find the region of visible light, X-ray, infra-red, microwaves, ultraviolet radiations, radio frequencies, gamma rays, cosmic rays.

An important aspect of spectroscopy is that a spectra of a compound (graph of the absorbed or transmitted radiation in function of the wavelength or the frequency) depends of its own structure. There are three types of absorbance spectroscopy most used: Infrared spectroscopy (IR), nuclear magnetic resonance spectroscopy (NMR) and ultraviolet-visible spectroscopy (UV-vis). Theorically these three differ only for the frequency of the radiation used. There is also another type of spectroscopy widely used that is mass spectroscopy (MS), but it is not an absorbance spectrometry.

Infrared radiation spans a section of the electromagnetic spectrum having wavenumbers from roughly 13,000 to 10 cm–1, or wavelengths from 0.78 to 1000 μm. It is bound by the red end of the visible region at high frequencies and the microwave region at low frequencies. The IR region is commonly divided into three smaller areas: near IR, mid IR, and far IR. We are going to focus our attention on mid IR region, between 4000 and 400 cm–1 wavenumbers that is also the most frequently used, but also the other regions are used for analysis (Sherman Hsu, Handbook of Instrumental Techniques for Analytical Chemistry). In IR spectroscopy, an organic molecule is exposed to infrared radiation. When the radiant energy matches the energy of a specific molecular vibration and a net change in the dipole moment of the molecule happens, absorption occurs. The IR radiation is absorbed and converted from the organic molecule in vibration energy. This absorption is quantized, but

the vibrational spectra appear like bands rather than lines and this is due by the fact that the changement in the vibrational energy is followed by variations in the rotational state. The bands are hence roto-vibrational bands. The main categories of molecular vibrations are stretching and bending. In the stretching type there is a change in inter-atomic distance along bond axes and we can have symmetric or asymmetric type of stretch. In the bending type there is a change in angle between two bonds, here there are four types of bend: rocking, scissoring, wagging and twisting. In addition to the vibrations mentioned above, interaction between vibrations can occur (*coupling*) if the vibrating bonds are joined to a single, central atom.

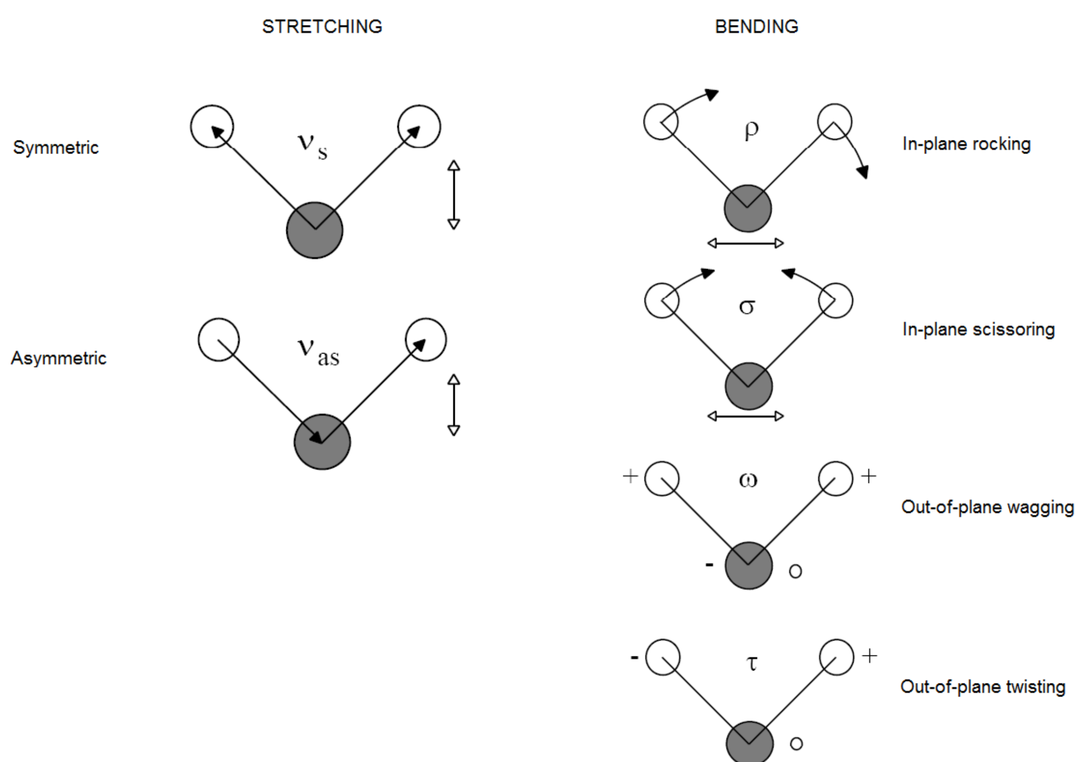

Figure 1: Different types of molecular vibrations (A. Piazzalunga).

Although fundamental absorptions or vibrations are visualized as strong bands or peaks, an IR spectrum is complicated by the presence of three types of bands: weak overtone, combination, and difference bands. Overtones result from molecules being excited from the ground state to higher-energy states and correspond to multiples of the frequency of the fundamental absorption. Combination bands result from frequencies that are vibrationally coupled, and difference bands result from the difference between two interacting bands. No two molecules of different structure have exactly the same IR absorption pattern; thus, the IR spectrum of a pure sample can be described as a molecular fingerprint characteristic of a specific chemical or biochemical substance (Esbensen *et al.*, 2008).

1.5.1 FTIR Spectroscopy

FTIR spectroscopy is a nondestructive technique that provides structural information on molecular features of a large range of compounds. Its main advantages are speed, high degree of automation, medium resolution, and cost-effectiveness. Recent improvements in instrumentation together with advances in fiber optics and chemometrics have provided an analytical tool that is suitable for routine qualitative analysis and process control in many industries (Esbensen *et al.*, 2008). Its basic principle is represented by the possibility to gather at the same time all the frequencies of the IR spectra on the detector. This is possible transforming, by an interferometer, the polychromatic IR radiation released by the radiation source (each instant with the same intensity) in an interferogram, where the absorption is not anymore in function of the frequency, but of the time (passing from the dominion of frequencies to the dominion of time). So, differently from the traditional spectrophotometers, in this equipment it is used the Michelson's interferometer which produces during a special scansion the interferogram of the examined compound. After the passage of the radiation through the sample, a mathematical operation known as Fourier transformation converts the interferogram (a time domain spectrum displaying intensity versus time within the mirror scan) to the final IR spectrum, which is the familiar frequency domain spectrum showing intensity versus frequency. This also explains how the term Fourier transform infrared spectrometry is created. In more technical view an incandescent light source emits a ray in the field of infrared; this ray arrives to a half-reflecting mirror (the beam splitter, a crystal of KBr) that divides the beam half to a fixed mirror and half to a moving mirror. The beams reflected by these two mirrors are sent back to the beam splitter that put back together the two beams and send them to the detector. These two beams made a different optical pathway that creates constructive or destructive interferences and creates a different signal proportional to the different optical pathway in the single moment. With this movement all the monochromatic radiations give a combined signal of interference (interferogram). Then the Fourier transformation gives a traditional IR spectra, transforming the signal in function of time to the signal in function of wavelength (Oliviero Rossi notes, University of Calabria).

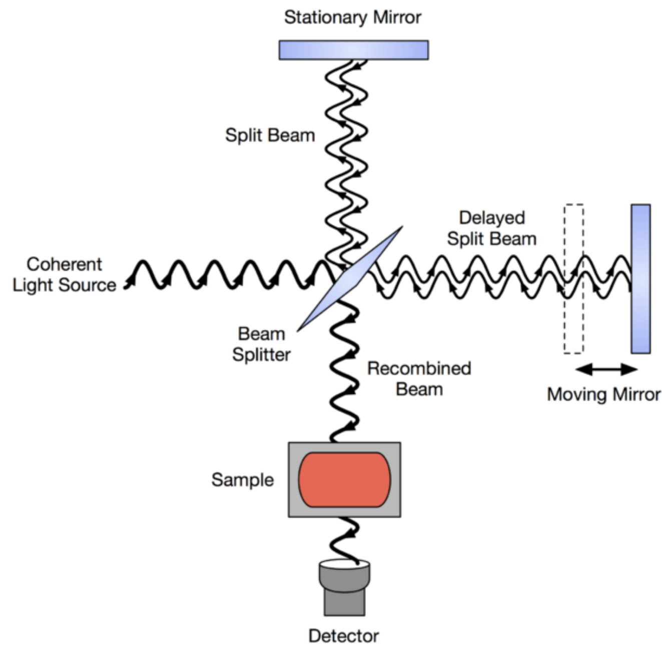

Figure 2: Schematic diagram of a Michelson interferometer, configured for FTIR. (www.en.wikipedia.org)

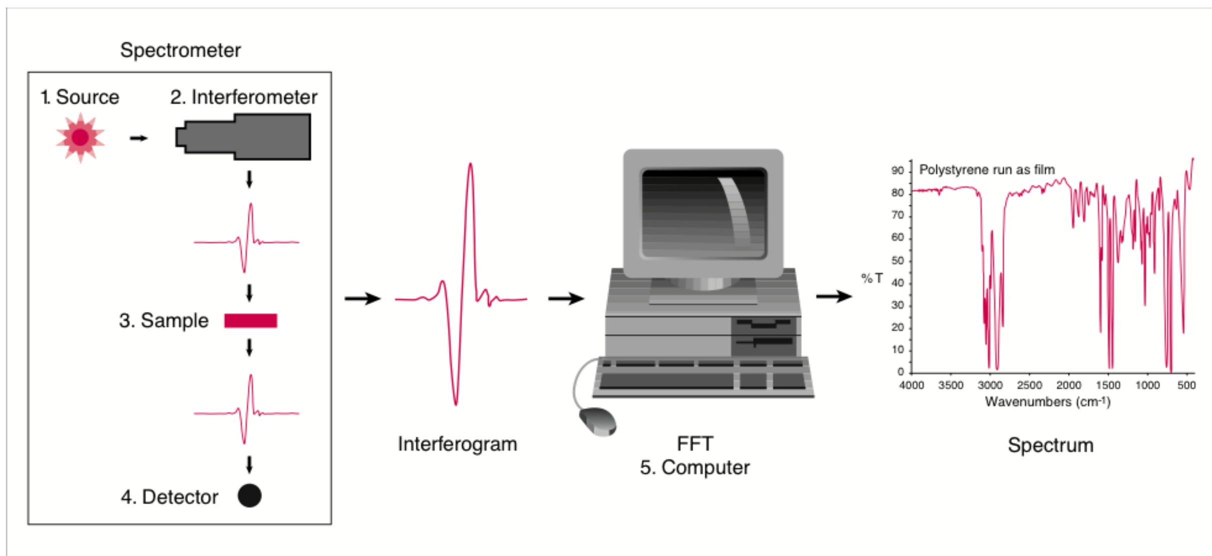

Figure 3: Data processing in FTIR.

(ThermoNicolet at *http://mmrc.caltech.edu/FTIR/FTIRintro.pdf*)

This technique in comparison with the traditional spectrometry offers some benefits, we are going to see the main of them:

1) Better speed and sensitivity. A complete spectrum can be obtained during a single scan of the moving mirror, while the detector observes all frequencies simultaneously.

2) The time of measurement is reduced to 1 sec or less versus 10 to 15 min of the conventional one.
3) Increased optical throughput. Energy-wasting slits are not required so the overall energy of the radiation source is available. More power arrives to the detector in comparison to the dispersive spectrometers.
4) Internal laser reference. The use of a helium laser as the internal reference provides an automatic calibration.
5) High precision of wavenumbers. It's possible to overlap at the signal as an internal standard the radiation of a laser source, which frequency is precisely known.
6) No heating of the sample.
7) Elimination of stray light and emission contributions.
8) Powerful data station.
9) Possibility to interface a gas chromatographic device.

1.5.2 Application of FTIR spectroscopy in enology

In the last decades FTIR spectroscopy developed really fast and was widely investigated and used in different fields of science. Several foods and beverages have been analyzed using FT-MIR and chemometrics (Briandet *et al.*, 1996; Budinova *et al.*, 1998; Cabezas *et al.*, 2006). In particular the world of enology showed a great interest in this application and used this equipment to obtain qualitative and quantitative information about the different parameters that are characterizing wine and must.

The use of FTIR in enology must be based on specific analytical calibrations established through the quantification by usual of physical and chemical methods of the parameters to analyse. The FTIR technique is an indirect analytical method (Patz *et al.*, 2000). The analytical calibrations for all the parameters are created after the quantification of those parameters in the calibration samples using routine methods and the analytical results are used as a reference for the FTIR equipment calibration.

In order to calibrate the equipment, it is necessary to measure the infra red spectrum of a group of wine samples and the parameters to calibrate should be quantified in those wines by routine methods. The selection of appropriate wave lengths for the quantification of each parameter is crucial in the development of new calibration, for which a great diversity of wines is required (Dubernet 2002; Moreira *et al.*, 2002).

The number of parameters of wine and must able to be detected by FTIR spectroscopy increased year by year from the development of this technique with the possibility now to determine parameters like ethanol, organic acids, density, total and volatile acidity, reducing sugars, dry extract and many others. Taken from the thesis of Ana Paula Nascimento of 2011, we can see table 1 as a resume of the application of FTIR spectrometry in the region of MIR in the enology field.

Table 1: Applications of FTIR spectrometry in the region of MIR in the viticulture-enology (Ana Paula Nascimento, 2011).

Applications	References
Characterization of polysaccharides in wine	Coimbra *et al.*, 2005; Boulet *et al.*, 2007
Determination of ethanol in wine	Gallignani *et al.*, 2005
Determination of organic acids in wine	Moreira et Santos, 2005
Estimation of procyanidins average degree of polymerization in grapes	Passos *et al.*, 2010
Determination of parameters of routine analyses in must and wine	Dubernet et Dubernet, 2000; Edelmann *et al.*, 2001 ; Moreira et Santos, 2004 ; Patz *et al.,*2004 ; Urtubia *et al.,*2004 ; Moreira et Santos 2005 ; Swanepoel *et al.,* 2007 ; Wynne *et al.,* 2007 ; Urtubia *et al.,* 2008 Fu *et al.,* 2009 ;
Determination of glycosidic precursors in grapes	Schneider *et al.,* 2004
Prediction of total antioxidant capacity of red wines	Versari *et al.,* 2010
Determination of glycerol in wines	Dixit *et al.,* 2005
Determination of total and volatile acidity in wines	Pizarro *et al.,* 2011
Determination of colored components in red wines	Versari, Boulton and Thorngate, 2004
Quantification of total tannins and average degree of polymerization in red wines	Fernandez and Agosin, 2007
Identification of spectral regions for the quantification of tannins in red wines	Jensen *et al.,* 2008
Determination of colloidal stability in wines	Versari *et al.,* 2011
Determination of assimilable Nitrogen in musts	Dubernet *et al.,* 2001; Skoutelas D., 2010

Considering the aim of this thesis we are going to examine more deeply just the quantification of tannins with FTIR spectroscopy. In particular we are going to analyze the different steps in the development of this spectroscopic methodology made by Fernandez and Agosin in 2007, Jensen *et al.* in 2008, Passos *et al.* in 2010 and Nascimento in 2011.

Wine characterization by FT-MIR spectroscopy presents two main limitations: similar IR absorption bands of most interesting compounds and dominating absorption of major wine components, particularly ethanol and water (Patz *et al.,* 2004). Both limitations are particularly critical for the analysis of phenolic compounds, because ethanol, water and organic acids absorb in the same MIR region, masking the characteristic IR vibrations of phenolics (Moreira and Santos, 2004).

Fernandez and Agosin investigate the suitability of FT-MIR spectroscopy as an accurate method for quantifying wine tannins and evaluated the prediction of the mean degree of polymerization. They developed the model on a single red wine variety, the samples were previously purified by solid phase extraction and as references were used the methods of protein precipitation (from Harbertson *et al.*) and phloroglucinolysis. Then several models were developed using multivariate PLS regression and spectral interval selection procedures. The main differences presented by the spectra were found in the wavelengths regions between 2800-3700 cm-1 and 800-1800 cm-1. The prominent absorption band around 3400 cm-1 can be associated with O-H and C-H stretching vibrations. The peaks of the fingerprint zone, between 800 cm-1 and 1800 cm-1 could be attributed to C=C-C aromatic ring stretching and several aromatic C-H out of plane and in plane bending vibrations. In particular, the peak around 1285 cm-1 was assigned to the ethereal C-O stretching vibration arising from the pyran-derived ring structure of flavanoid-based tannins. As a whole, this work showed that FT-MIR spectrometry combined with multivariate data analysis allows an accurate determination of tannin concentration and mDP in wines.

Also in Jensen *et al.* study of direct quantification of tannins in wines by infrared spectroscopy is faced the problem of the interference between the characteristic absorption bands of major wine components and tannins. This problem has been overcome by sample purification using solid phase extraction (as previously shown in the study of Fernandez and Agosin) but it is still a not suitable strategy for rapid tannin analyses in wine industry. So Jensen study focus on an alternative way, which is to identify the characteristic spectral regions of tannins, which do not suffer from this interference, and in turn use this identification to develop calibration models that then allow the rapid quantitative assessment of tannins by FT-MIR. So a four variable selection methods for finding the best regions in the range of the IR spectra were evaluated to see if the calibration models could be improved. Although the regions identified by the four variable selection methods were not identical, two regions were selected by all four methods: the region from 1060 to 995 cm-1, which was dominated by high absorption of the OH stretch in ethanol, and the region between 1485 and 1425 cm-1, at which grape tannin gave a distinct absorption peak. The present study

demonstrated that the identified regions could be used to develop calibration models, which allowed the measurement of tannins in wines by FT-MIR spectroscopy.

In the work of Passos *et al.,* regression models based on the FTIR spectral region between 1800 and 700cm−1 using Partial Least Square regression and orthogonal projection to latent structures were assayed in grape seeds freeze-dried procyanidin extracts. Identification of procyanidins molecular features related to the absorbance characteristics in the mid infrared region were established and correlated with procyanidins average degree of polymerization (DPn). For this purpose it was used a methodology of fractionation based on the difference in the solubility of procyanidins in methanol/chloroform solutions, showing the relationship of procyanidins DPn and solubility in this solution. This method allowed also to build the calibration model in order to estimate the average degree of polymerization of grape seed procyanidins. According to the regression model it is possible to identify the most important bands related to the DPn value. The observation of the positive peaks at 1203 and 1099cm−1 suggests a possibility to assess the DPn of procyanidins by monitoring the aromatic substitutions, in this case the increase of aromatic substitutions in polymerized molecules. Although this method has not been yet tested in other than grape seeds it is expected that applications can be found also in other procyanidins sources like grapes extract or wine.

The recent study of Ana Paula Nascimento has the objective to evaluate the application of FTIR spectroscopy for the quantification of total condensed tannins, as the determination of their three fractions: polymeric, oligomeric and monomeric. The calibration models were developed using two methodologies as reference: fractionation by reverse phase method and quantification by reaction with vanillin and the method of precipitation with methyl cellulose. Unfortunately the results leads to the conclusion that in this study the models are not robust enough for the estimation of concentrations of total condensed tannins and the three fractions. With this last survey it is clear that the problems came out in the previous studies are still not easy solved, especially when the spectroscopic analysis is directed directly to the raw wine. So we can understand that the important goal to find a rapid and precise analytical tool for the quantification of total tannins with FTIR spectroscopy is far from been achieved, despite the elevate number of studies in relation to this topic.

1.6 Chemometrics

With the introduction of modern analytical instrumentation capable of multivariate responses like FTIR spectroscopy, we can have in an easy and rapid way indirect, full-spectrum measurements of the chemical and physical properties of a sample. As a result, extremely large data sets are generated in which essential information may not be readily evident. In fact we need more complex analytical approaches than the traditional univariate methods. For this purpose, chemometrics and multivariate analysis were joined closely to the quantitative analysis in the infrared spectrometry. Now the use of chemometrics for calibration, validation and comparison of main components of wine as determined by IR is indispensable. Qualitative and quantitative IR spectroscopic methods typically require multivariate calibration algorithms to model the spectral response to chemical or physical properties of the calibration samples. Measured frequencies often have to be preprocessed through a series of mathematical procedures, including various forms of scaling and corrections to yield a suitable absorbance spectrum. Vibrational bands (peaks) are generally overlapping and may often appear nonspecific and poorly resolved (Bauer *et al.,* 2008).

In particular with the term chemometrics is considered the use of mathematical and statistical techniques for handling, interpreting, modelling and predicting a group of data. Basically is a science that uses mathematic-statistics methods in order to solve multivariate problems. The birth of this discipline is brought back up Svante Wold and Bruce Kowalski that in June 1974, in a letter sent to Analytical Chemistry journal, first suggested the term "Chemometrics" (D'Angelo *et al.*). Chemometrics permits a multivariate approach to the system under study: in this way it permits to take in account all the variables, allowing the best use of the information in the analysed dataset. Nowadays chemometrics regroups the methods of classification, modelling and multivariate regression, the analysis of similarity, the analysis of principal components and the different methods linked to it, the expert systems and the methods of artificial intelligence, strategies based on the neural network, methods of experimental design and optimization (www.chemiometria.it). Chemometrics pursues three principals objectives:

a) Organize an optimal sampling (representative);
b) Optimization of the analytical methodologies;
c) Extract the maximal useful information from the data obtained (Sinelli N.).

1.6.1 Multivariate Analysis

In chemometrics a lot of emphasis is given to the multivariate analysis of the problem. In fact nowadays analytical instruments like FT-IR spectroscopy gives to the researcher a dataset, in the specific case a spectra, that must be approached as a multivariate data in order to gain the maximum of information from it. All these data values are difficult to cope and evaluate in a simple and fast way using simple univariate statistical tools, specially due to their large number and to their multivariate correlation. In order to discover relevant patterns within large multivariate data sets, the application of modern chemometrics methods based in statistical multivariate data analysis is proposed. The basic assumption of chemometrics methods is that each of the measured parameter in a particular sample is affected by contributions coming from multiple independent sources (Tauler R. *et al.*). As pointed out by Rasmus Bro in 2003 multivariate data analysis and multivariate calibration give several significant advantages in their application. It is claimed there, that multivariate models are more adequate than univariate models. This strong statement holds in general because it is always possible to discard variables such that a univariate approach is re-obtained and multivariate models can handle situations that cannot be handled univariately. In particular, it is possible to minimize the influence of the noise, to incorporate interferents and to have automatic outlier detection when building or using a model. In addition multivariate models and data make it possible to supplement the traditional deductive approach with an exploratory one (Rasmus Bro, 2003). If multivariate data analysis is not used we can encounter some risks like spurious correlations, lost of relevant information and not recognizable outliers in a simple comparison.

1.6.2 Data pre-treatment

In general the dataset of information is in form of matrix where each row is an object, described by many variables one for each column. The instrumental data gives an high information but correlated at the analytical data in a complex way. There is as well an elevated noise and not relevant information. In order to remove or reduce the sources of systematic or unpredictable variation which are not related with the chemical problem of interest, as well as to increase the spectral characteristics, a pre treatment of the data is necessary (A. P. Nascimento, 2011). Pretreatments are performed with very simple

mathematics, but it is very important to understand perfectly what they are doing on the dataset.

In the Infrared field the selection of the type of pretreatment of the data depends on the aim of the analysis and on the type of data. In general the most diffused mechanisms of normalization are the followings: Normalization, Offset and Multiplicative Scattering Correction, Derivatives processes. In addition to these, data can undergo to procedures of centering and scaling. For a better understanding we are going through these mechanism in the specific.

Normalization

Normalization procedure assumes a relevant importance in the phase of data pre-processing. It is necessary so that important process variables with low amplitude are not considered less important than variables with more huge amplitude. There are different methods of normalization like: min-max normalization, Z-score normalization, robust normalization, vectorial normalization. In spectroscopic applications, scaling differences arise from pathlength effects, scattering effects, source or detector variations, or other general instrumental sensitivity effects. The sample normalization preprocessing methods attempt to correct for these kinds of effects by identifying some aspects of each sample which should be essentially constant from one sample to the next, and correcting the scaling of all variables based on this characteristic. Normalization also helps to give all samples an equal impact on the model. Typically, normalization should be performed before any centering or scaling or other column-wise preprocessing steps and after baseline or offset removal (www.wiki.eigenvector.com).

Here we have an example of how a normalization methods works with spectra, taken from a research of Martens, Nielsen and Engelsen in 2003. These spectra were measured as 20 replicates of 5 synthetic mixtures of gluten and starch. In the original data (top plot), the five concentrations of gluten and starch are not discernable because of multiplicative and baseline effects among the 20 replicate measurements of each mixture. After normalization (bottom plot), the five mixtures are clearly observed in groups of 20 replicate measurements each.

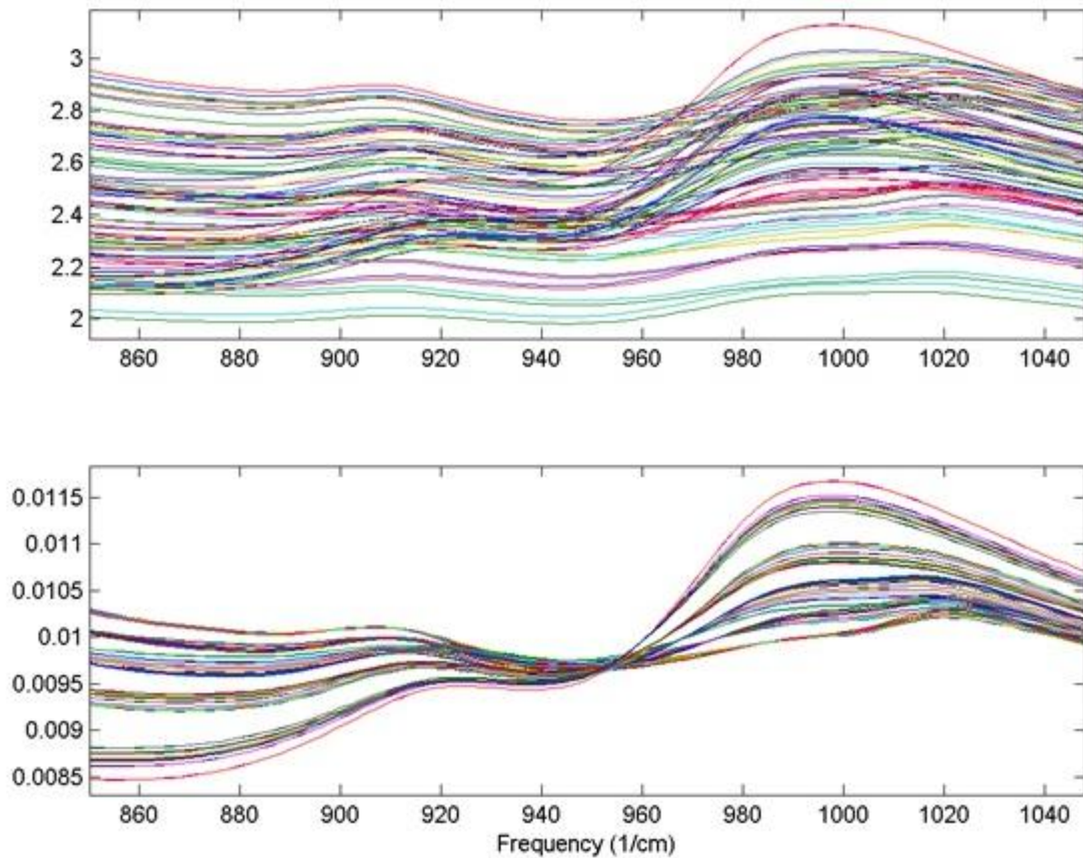

Figure 4: Effect of normalization on near-IR spectra of five synthetic gluten and starch mixtures. Original spectra (top plot) and spectra after normalization (bottom plot) are shown (Martens, Nielsen and Engelsen, 2003).

Multiplicative Scatter Correction (MSC)

Offset correction is an old method applied to correct flat background by Candolfi *et al.* in 1999.

Multiplicative scatter correction (MSC) is a relatively simple processing step that attempts to account for scaling effects and offset (baseline) effects (Martens and Næs, 1989). Basically it is a normalization of the spectra which removes the effects of the variability between samples due to the difference in the optical pathway. The correction of scattering provides a linear transformation of all the spectra in a way that every single spectra is going to be as similar as possible to the average spectra of the dataset. This correction is achieved by regressing a measured spectrum against a reference spectrum and then correcting the measured spectrum using the slope (and possibly intercept) of this fit.

An extension of the MSC correction is the extended multiplicative scatter correction (EMSC). It keeps into account the variation of the wavelength linked to the physical or chemical

phenomenon that can influence the spectra. By the use of the knowledge of the phenomenon that influence the spectra, can be obtained a good correction of the scatter's phenomenon.

Derivatives

Another very common method for background correction is mathematical derivative of the spectrum (Tahboub and Pardue, 1985) usually first derivative and second derivative. The calculation of the derivative is a technique widespread in the IR applications. The information inside a spectra can be better analyzed when we are working with derivatives. A flat background can be removed by first derivative, and a sloping background can be removed by second derivative. In fact, derivative can be seen as removing low-frequency components and amplifying high-frequency components of the spectrum. In particular second derivative measures the changement in the slopes of the curve, removes the effects of backgrounds due to the scattering, helps in the resolution of overlapping peaks and maintain the same position of the peaks of the original spectra. By selecting proper order of derivative, the background can be removed as low-frequency components, but meanwhile the noise is amplified (Hai-bin Qu *et al.,* 2005).

1.6.3 Principal chemometrics techniques

In the following pages we are going to briefly see in details the main techniques used within chemometrics, in particular regarding the explorative analysis of the dataset and the methods of multivariate regression. Not using these methods could cause the loss of part of information that can be extracted by the dataset, because are not considered the multivariate, synergic and antagonistic effects, that can be put in evidence only by multivariate techniques (chemiometria.it).

1.6.4 Explorative analysis

Explorative analysis permits to detect and evaluate the structures and connections also inside very complex dataset, collecting as much information as possible concerning the problem in exam.

A basilar technique is Principal Component Analysis (PCA) that has got the target to extract the maximum information possible from a multivariate structure of data, synthesizing it in few linear combination of the variables themselves. This method gives a general vision of the problem, the relationship between the objects or the classes and a preliminary indication on the role of the variables, highlighting the possibility of eliminating the ones that are bringing similar information. The simplification of this huge amount of parameters, necessary to have a better sight of the data, is achieved by rotation. In fact it consists in a process of rotation of the original data, carried out in such a way that the new first axis (that is going to constitute the first principal component) will be oriented in the direction of maximum variance of the data, the second axis will be orthogonal at the first and will be in the following maximum variance of the data, and so on for all the new axis. The number of these new axis (the principal components) is going to be equal to the number of the original variables. Because the principal components are the axis relative to the direction of maximum variance in a decreasing order, the first principal component is going to explain more percentage of the variance, the second a bit less, and so on until the last components. In these cases is possible to erase part of the residual variability, taking in account just a small number of components. Another important aspect of PCA is that it gives an algebraic solution with very interesting possibilities of graphic representations.

In conclusion we can say that PCA is very useful and provides a compression of the dataset and a reduction of the noise, with the only drawback that a change in the dataset modifies all the factors.

1.6.5 Multivariate regression

The multivariate regression methods are used to make quantitative predictions relatively to one or more properties of the system studied. Their aim is to find the better relationship between a group of variables that describes the objects studied and a group of measured answers from the same objects. In particular in the infrared spectrometry the algorithms of regression provide a method to correlate the spectral information with the concentration of the constituent products. The principal models of calibration can be divided in four different typologies, in order of complexity: linear regression or ordinary least squares (OLS), multiple linear regression (MLR), principal components regression (PCR) and partial least squares regression (PLS).

Linear regression

In the methods of linear regression, a single point of the spectra (or a peak) is correlated with the parameter to predict in order to generate a linear equation.

Multiple linear regression

This type of regression is a simple variant of the linear regression in which the points correlated with the predictors are not only one but numerous.

Principal components regression

PCR combines a PCA with a multiple linear regression to create a quantitative model. PCA is made to define the number of factors to study. This means that, at the beginning, are calculated the principal components of the matrix of the predictors (matrix of spectral data) and that are used only the first principal components, so that ones that contain the maximum information. PCR uses only the relevant information contained in the matrix of spectral data not considering the one with the value of concentration. In this way is possible to reduce a lot the background noise. On the principal components so extracted is carried out the regression with the objective to obtain a calibration or predictive model.

Partial Least square regression

PLS is an upgrade of PCR, because the used components are derivate not only from the spectral dataset, but also from the matrix with the value of concentration. In this way it is possible to maximize the variance of the two matrix of the system. Doing so the choice of the principal components, that are used for the regression is done in a more accurate and efficient way. PLS so uses better the information, reducing the effect of huge but not relevant variation in the predictors. Partial least square regression shows a certain flexibility that permits to adapt this method to situation where the use of traditional multivariate methods is very limited. Moreover, PLS can be used as an instrument of explorative analysis in the selection of satisfying predictors and to identify outliers before the classical linear regression. In general PLS gives simpler model than those built by PCR and is able to give good answers also in presence of data not so much precise.

2. AIM OF THE THESIS

Grape-derived condensed tannins, also known as proanthocyanidins, are critical quality components of red wine. The aim of this thesis is to investigate the suitability of Fourier transform mid-infrared spectroscopy and chemometrics techniques for the accurate quantification of red wine condensed tannins. As a reference for the calibration of the models, it was used the tannin quantification by methylcellulose precipitation. In order to have a better understanding of the feasibility of the fractionation methodology proposed by Sun *et al.* in 1998, it was also performed a single-laboratory method validation for the fractionation of wine proanthocyanidins according to their degree of polymerization by C_{18} Sep-Pak cartridges and their quantification by vanillin reaction.

3. MATERIALS AND METHODS

This chapter explains the operational part of the thesis and the methodologies that stands at the basis of this study. For a better comprehension it is divided in two parts.

In the first part are grouped the chemical methods used for the research. So it is described the methodology of reference developed by Sarneckis *et al.* (2006), that is based on a measure of 280 nm absorbance of phenolics, before and after precipitation of tannin by Methylcellulose. This accounts for the 280 nm absorbance of tannin. The other methodology described is the one developed by Sun *et al.* (1998a, 1998b) for the fractionation in reverse phase of the three condensed tannins fraction FI, FII and FIII of red wines, followed by its quantification by colorimetric reaction with vanillin. As well are described in this part the basis of the single laboratory validation of this last methodology.

In the second part is described the FTIR methodology applied, in particular the spectral regions used for the calibration models, the conditions of acquisition of the spectra and the statistical approach followed.

3.1 Sample characterization

In this study, in order to have a better evaluation of the FTIR method of analysis with the Methylcellulose method of reference, it was chosen to provide a wide group of red wine sample. This with the aim to provide a good heterogeneity of the condensed tannin values within the samples. So it has been chosen wines with different ages and different varieties.

In total were analyzed 88 samples of wines ranging from the oldest year 1969 to the year 2011. The wines were all coming from *Vitis vinifera* grapes and from different varieties mainly: Castelao, Alicante, Pinot Noir, Merlot, Alfrocheiro, Aragones, Syrah, Cabernet Franc and Sauvignon, Touriga Nacional, Trincadeira, Ramiscos and Touriga Leira. The wines had total tannins concentrations that were distributed between 301-3577 mg/L catechin equivalents.

In order to have a laboratory validation of the fractionation method developed by Sun *et al.*, it has been chosen, between the wines analyzed previously, one sample with an average concentration of total condensed tannins. The aim of this choice was to avoid extreme cases

of concentration of total condensed tannins in the wine that were not considered significant for the validation process.

3.2 Chemical methods

3.2.1 Quantification of total condensed tannins by methylcellulose precipitation

The methodology followed is based on the method developed by Sarneckis *et al.* in 2006.

Before proceeding two solution were prepared:

- Solution of satureted ammonium sulphate

In a flask with 300 mL of deionized water it was added ammonium sulphate in crystal form, stirring until it has been reached a deposit, in the bottom part of the bowl, of around 1,5 cm of ammonium sulphate crystals. This to make sure that the solution is well satureted.

- Solution of methylcellulose 0,04%

In a bowl of 1L containing 300 mL of deionized water at 80°C, it was added 0,4 g of Methylcellulose (M-0387, Sigma-Aldrich®, 2% of viscosity in a water solution at 20°C) and quickly dissolved with a magnetic stirrer, avoiding in this way the development of big gelatinous agglomerates. After this, it was added 700 mL of deionized water at 0-5°C and it was stirred for 20-40 minutes submerged in ice water (0-5°C). So it was taken out from the ice water and after the solution reached the ambient temperature, the final volume was adjusted to 1L with deionized water.

The procedure is based on the addiction of 3 mL of Methylcellulose 0,04% (v/v) in 0,1 mL of wine, stir it several times and leave the solution for 2-3 min. (The choice of the addition of 0,1 mL of wine, and the subsequent dilution, was taken after some trials of the condensed tannin samples range and according to the fact that by allowing a 5% underestimation of the tannin concentration it was found that the tannin response should lie between 0,3 and 0,75 absorbance units) (Jensen *et al.,* 2008). Then was added 2 mL volume of satured solution of ammonium sulphate and the volume was adjusted until 10 mL with deionized water. The solution was left still for 10 minutes and then was centrifuged for 5 minutes at 4.000 rpm. A control solution was also prepared for every sample, in this solution the procedure was the same as above but without the addition of methylcellulose solution. Just after the centrifugation the surnatant was transferred in a quartz cell and measured its absorbance at 280 nm with a spectrophotometer Unican UV/Vis Spectrometer UV4 model. The final

determination of the total condensed tannins in the sample was obtained by the difference of the absorbance values of the solution with methylcellulose and the control solution. So:

A_{280} (total condensed tannins) = A_{280} (control sample) – A_{280} (sample solution)

From this value of absorbance it is possible to obtain the equivalent in mg/L of catechin through a calibration curve: absorbance 280 versus catechin. This curve was previously obtained by correlating the variation of concentration from 10 to 100 mg/L of catechin in water and the relative values of absorbance, with a R^2 value of the curve higher than 0,99.

The final concentration of total condensed tannins was corrected by a dilution factor:

Total condensed tannins (mg/L catechin Eq.) = tannins (mg/L cat. Eq.) x F(dilution factor)

In our case the dilution factor was 100. (0,1 mL of wine in 10 mL total volume)

3.2.2 Fractionation of proanthocyanidins by solid phase extraction

First step was the activation of the cartridges C_{18} Sep-Pak®Plus connected in series: the superior one is tC18 Sep-Pak®Plus and the inferior is C18 Sep-Pak®Plus. Activation process consists in the successive elution through the cartridge of 10 mL methanol, two elution of 10 mL each of distilled water and finally 10 mL phosphate buffer solution (pH 7,0). In the meanwhile 3 mL of wine sample was dealcoholized by rotary evaporation under vacuum Rotovapor R-200 (temperature should be maintained under 30°C during this operation) and adjusted to pH 7,0 with NaOH (0,1 M) solution and 20 mL phosphate buffer solution. This sample was then introduced mL per mL into the cartridge and then elution was carried out by 10 mL of phosphate buffer solution pH 7,0 diluted in distilled water in 1:8 proportion. This first elution permits to eliminate phenol acids from the sample. After the cartridges were dried with Nitrogen stream, elutions were carried out first with 25 mL of ethyl acetate to elute catechins and oligomeric PA, accompanied by some other small phenolic molecules (FI and FII fractions), and then with 15 mL of methanol to elute the polymeric PA (FIII fraction). These fraction were evaporated and the cartridge was activated as previously described. For the separation of the two fractions FI and FII the evaporated sample containing these two was dissolved with 3 mL of phosphate buffer solution pH 7,0 and introduced mL per mL into the cartridge.

After the cartridges were dried with Nitrogen stream, separation of catechins and oligomeric PA was realized by sequential elution with 25 mL of diethyl ether (FI) and then with 15 mL of methanol (FII). These two fractions were suddenly evaporated in order to eliminate the eluent solution.

The cartridge was used until three times, after a cleaning elution each time by 10 mL solution HCl 1% (v/v) in methanol and 10 mL methanol.

Fractionation procedure can be schematized as shown below the figure 5.

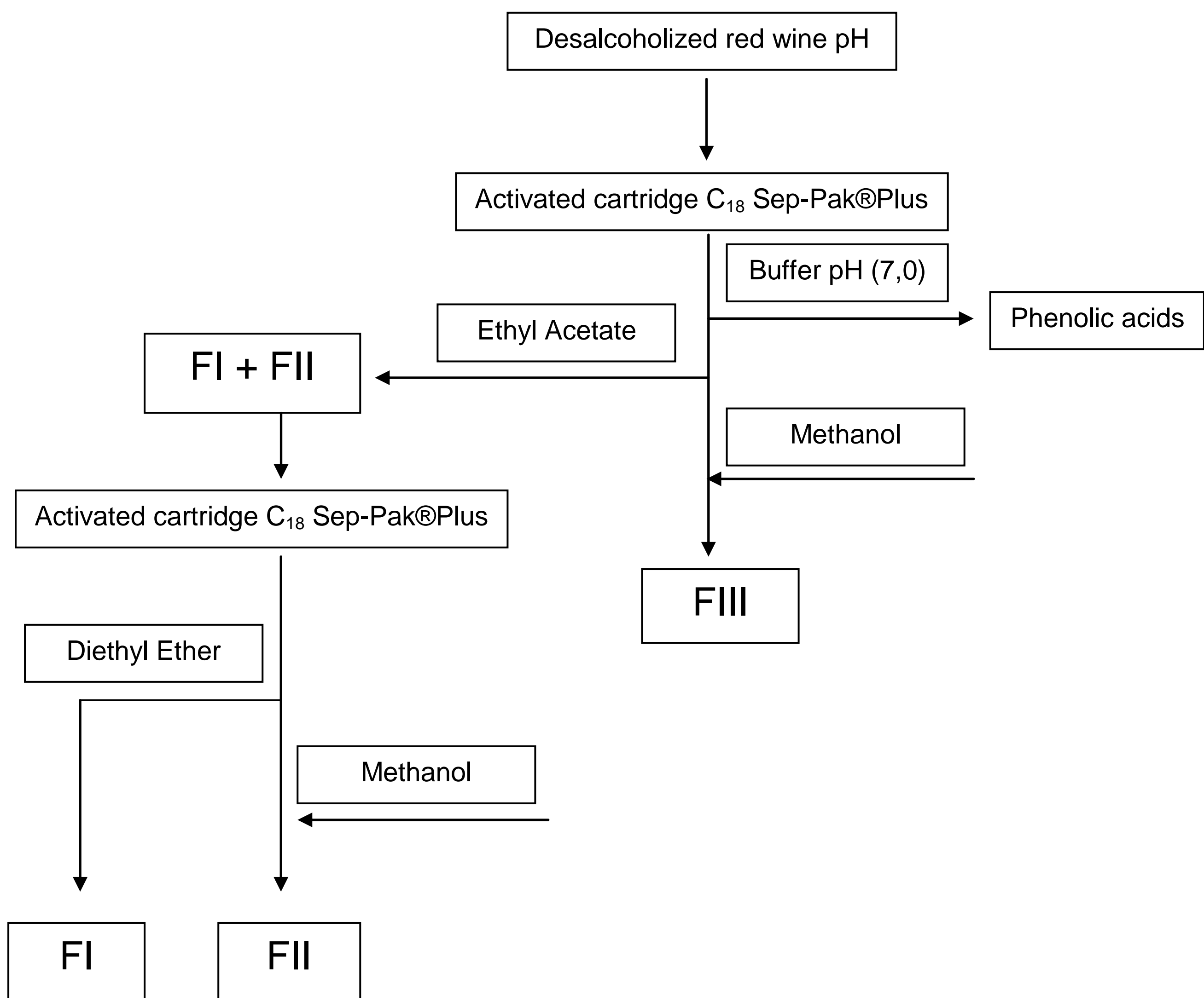

Figure 5: Fractionation procedure of proanthocyanidins by solid phase extraction.

3.2.3 Quantification of FI, FII and FIII fractions by vanillin reaction

The three fractions (FI, FII and FIII) coming from the separation through C_{18} Sep-Pak®Plus cartridges and containing respectively monomeric catechins, oligomeric PA and polymeric PA, were evaporated to dryness under vacuum at 30°C. Then the residue was dissolved in methanol in order to give a desired concentration for the quantification.

In a test tube, one milliliter of this methanolic solution was mixed first with 2,5 mL of vanillin solution (concentration of 10g/L in methanol) and then with 2,5 mL of 25% (v/v) H2SO4 in methanol; followed by homogenization to undergo vanillin reaction. The blank was simultaneously prepared in the same way except that vanillin solution was substituted by methanol.

The vanillin reaction with catechin fraction (FI) was carried out in a 30 °C water bath for 15 min, and the measurement of the absorbance at 500nm was also performed at a temperature between 25 and 30 °C with a spectrophotometer Unican UV/Vis Spectrometer UV4 model. For oligomeric PA and polymeric PA fractions (FII and FIII), both the vanillin reaction and the measurement of *A*500 were performed at room temperature and the maximum *A*500 was taken as the measured value. Three standard curves, with R^2 higher than 0,99, should be conducted using (+)-catechin, purified oligomeric PA, and purified polymeric PA, to express, respectively, catechin content, oligomeric PA content, and polymeric PA content in the wine samples. From these curves it would be possible to calculate the quantity of each fractions in the sample expressed in mg/L of epicathechin.

3.2.4 Validation of analytical method

In order to support the MCP method and have more detailed vision of condensed tannin part with FTIR spectroscopy, the methodology of the fractionation in reverse phase of Proanthocianidins using C_{18} Sep-Pak®Plus cartridges was tested. After some weeks of trials this methodology was shown as time consuming and not precise as expected in order to prove the FTIR technology. So it was decided to do a validation of the method for this laboratory as it was done previously for the Methylcellulose precipitation method.

Method validation makes use of a set of tests that both test any assumptions on which the analytical method is based and establish and document the performance characteristics of a

method, thereby demonstrating whether the method is fit for a particular analytical purpose. Typical performance characteristics of analytical methods are: applicability, selectivity, calibration, trueness, precision, recovery, operating range, limit of quantification, limit of detection, sensitivity, and ruggedness (Thompson *et al.,* 2002).

For this single-laboratory method validation were chosen to be evaluated the following characteristics: precision and spike recovery.

Precision

It refers to the degree of accordance between independent results obtained by an analysis procedure in well defined conditions. These conditions are identified by some factors: method of analysis, material examined, laboratory, operator, equipment, calibration of the equipment, time between the analysis. Considering that the method and the material examined must be always the same, there can, be depending on the variation of the other factors, two types of precision: repeatability and reproducibility.

Repeatability is defined as the precision obtained in the conditions that all the previous factors are constant (same method on identical test material under the same conditions: same operators, same apparatus, same laboratory and a short period of time).

Reproducibility is the precision obtained using the same method on identical test material, under different conditions (different operators, different apparatus and/or different laboratories and/or different time).

With the fractionation in reverse phase and reaction with vanillin it was not possible to maintain all the conditions constant and so calculate the repeatability value. In fact for this methodology was not possible to analyze the same sample ten times in the same day, because the time for a single analysis was too long and it needed a huge quantity of laboratory equipment.

Thus the reproducibility was calculated using the same operator, the same apparatus, the same laboratory but in different days. In this way the red wine sample was analyzed ten times and each time the values of tannins of the three fractions FI, FII and FIII were recorded.

Spike recovery

Spike and recovery of known amounts of analyte into the various sample is a critical experiment to validate the accuracy of a given method.

It is defined as the recovery of analyte spike added to sample prior to sample preparation. The determination of spike recovery is based on results provided by spiked and unspiked sample and it is used to estimate matrix effects and sample preparation losses. In our case it was possible to add catechin as analyte spike at the concentration of 5 mg/L.

3.3 Fourier Transform Infra-Red methodology

3.3.1 Spectral acquisition of the samples

Fourier transform infrared analysis (FT-IR) was conducted by using multispec equipment, FT-IR AVATAR 380 NICOLET (Thermo Nicolet Corporation, Madison, WI) equipped with a transmission flow-cell (including Nicolet EZ Omnic FT-IR software package). The spectral acquisition was performed directly on the red wine samples without any pre-treatments. For the same sample two spectral acquisition were done. Bacchus Acquisition software was used to define measurement parameters.

The quantification of total tannins by Methylcellulose precipitation was performed in the same day as the spectral acquisition in order to avoid chemical alteration of the samples due to long time between analysis and so try to minimize the source of errors introduced in the dataset. Software OMNIC, Thermo Nicolet was used for the spectra acquisition. The data were storage with the whole spectral range, between 4000-500 cm^{-1}.

3.3.2 Development of calibration models

Considering the demonstrated importance of optimization of wavelength range, different strategies were tested in order to select the optimal wavelength range. (Williams and Sobering, 1993; Jensen *et al.,* 2008). So the models were built based in the following

spectral regions: full range region, main frequencies region and a region corresponding to the one suggested by partial least square regression with Optimize software.

For the main frequencies region, according to the studies of Fernandez and Agosin (2007) and Nascimento (2011), in order to build the calibration models just the most informative regions in the spectra were used. So in this case only five spectral regions were considered: 1000-1225, 1230-1320, 1450-1510, 1580-1615, 1660-2000 cm^{-1}.

In this region called also fingerprint zone (800-1800 cm^{-1}) were isolated these range of frequencies attributed to C=C-C aromatic ring stretching, several aromatic C-H bending vibrations and a characteristic peak assigned to the ethereal C-O stretching vibration arising from the pyran-derived ring structure of this class of tannins.

For the last region, it was used an automated optimization that checks common frequency regions in combination with several preprocessing using the software Optimize, an option of the Quant 2 package of Brucker OPUS. The regions suggested by the software optimization was these two intervals of wavelengths: 1446,42-1511,99 cm^{-1} and 2846,6-3062,6 cm^{-1}.

The models for the linear calibration by partial least square regression were developed for the quantification of total condensed tannins, using the reference method of Methylcellulose precipitation. For the calibration, partial least squares (PLS) regression was used, among the most commonly employed multivariate calibration method, for the evaluation of Mid infrared spectra. Using the cross-validation the same samples are used for both calibration of the model and testing.

Before the development of the calibration some outliers were taken out from the dataset because they showed problems in the spectral acquisition like the presence of abnormal peaks due to interferences in the acquisition or because they presented abnormal value due to problems during the quantification with the method of reference.

Prior to calibration the outliers were removed according to OPUS Quant software and different preprocessing treatments of the data matrix were tested, taking only the best combination of the treatments. This decision was based on the best coefficient of determination (R^2) and the best residual prediction deviation (RPD).

For the development of the calibration models for the quantification of tannins by FT-MIR spectroscopy using the Methylcellulose precipitation data as reference, samples were split into a calibration set (67 wines) and a validation set (76 wines) with similar standard deviations and comparable ranges of the tannin levels. This division choice was taken to

have a good calibration of the data but also an accurate error estimation, in order to robust testing the calibration model.

The calibrations models were developed by partial least squares regression (PLS-R) using the Bruker OPUS Quant software package vs. 7.0 (Bruker Optics, Ettlingen, Germany) using an automated optimization that checks common frequency regions in combination with several preprocessing using the software Optimize an option of the Quant 2 package. The number of PLS components (rank) was obtained by means of full cross-validation with 4 samples omitted. All models were calculated to a maximum rank of 15 and the results of the cross-validation [coefficient of determination (R^2), root mean square error of cross-validation (RMSECV) and residual prediction deviation (RPD)] and the test-set-validation [R^2, root mean square error of prediction (RMSEP) and RPD] were compared.

RPD value is the ratio of the standard error of prediction (SEP) to the standard deviation of the original data. The RPD should be as high as possible. A value of greater than 10 is excellent and equivalent, or better, than reference chemical or physicochemical reference method. Values of 5-10 are adequate for quality control, and 2.5 and over are satisfactory for screening. On the other hand an RPD value of 1.0 means that the SEP and the SD are the same and the instrument is not capable of predicting the parameter accurately, using that calibration.

R^2, SEP, RMSEP and RMSECV are measures of a model's fit to the data and of predictive power used in our study. In the measures considered, we are attempting to estimate the average deviation of the model from the data.

There is variance in the known concentration values, R^2 tells how much of that variance is reproduced by the predicted values. Higher values proximal to 1 indicate a better correlation. Considering just the value R^2 it’s not the best choice for measuring the merit of a calibration, other measure are introduced. Standard error of prediction (SEP) is a measure of the variability of the difference between the predicted and the reference values for a set of validation samples.

RMSEP and RMSECV measures the precision of the test analysis. These are calculated using the same formula, except that in RMSECV the predicted concentration are from cross validation and in RMSEP the predicted concentrations are from calculated from the test samples.

The root-mean-square error of cross-validation (RMSECV), is a measure of a model's ability to predict new samples. It can only be calculated if the model has been cross-validated. It should be evident from the comparison of the models that fit and prediction are entirely

different aspects of a model's performance. If prediction is the goal, a model that gives the minimum RMSECV value amongst the prediction models should be selected. For comparison of models intended for prediction it is inadequate to look just at model fit. For PLS, the highest predictive ability with the smallest number of factors is advantageous in that there are fewer factors to interpret. (Ozgur Yeniay and Atilla Goktas, 2002; Davies and Fearn, 2006).

4. RESULTS AND DISCUSSION

4.1 Quantification of total condensed tannins by Fourier Transform Infrared Spectroscopy

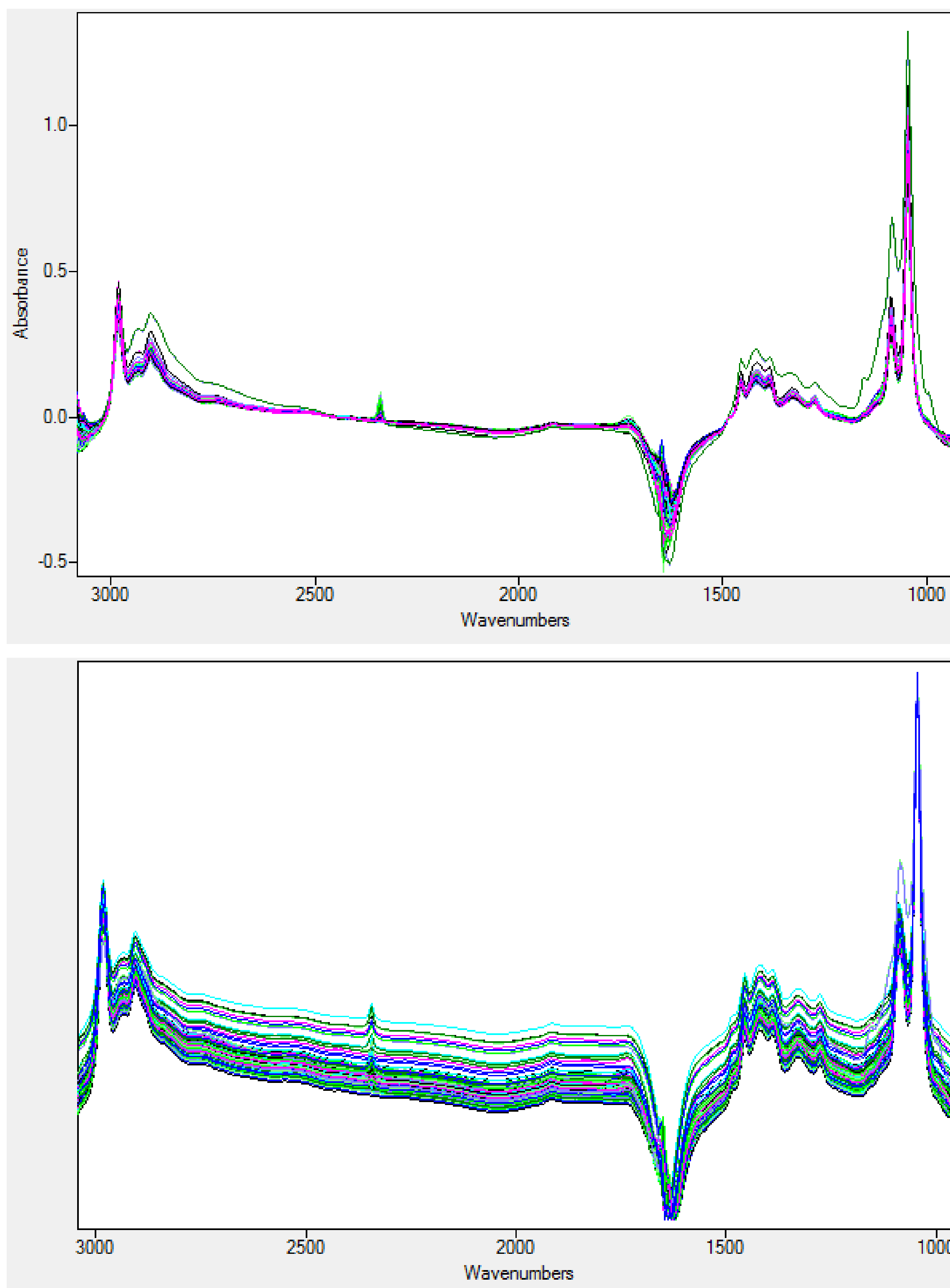

Figure 6-7: Red wines spectra in FTIR (3100-900 cm^{-1}) in overlay (6) and superimposed mode (7) (taken with Essential FTIR v. 3.0, Operant LLC).

In this study the spectrum of 153 red wine samples was used for the analysis with Fourier Transform Mid Infrared Spectroscopy for the quantification of total condensed tannins. In figures 6 and 7 we can see the spectra in overlay mode (6) and superimposed mode (7). Overlay mode results in spectra being plotted on common Wavenumbers (3100-900 cm^{-1}) and Absorbance (-0.5 and 1.5 A) axes hence they are overlaid, superimpose mode matches the y axis magnitude of all data. This last feature is useful for comparing spectra collected at different path lengths or concentrations. Notice that the visible absorbance features are re-scaled so as to match each other magnitude as closely as possible. No tick-marks are placed on the y-axis because they may be misinterpreted as they do not apply to all the spectra in the display (Essential FTIR v. 3.0, Operant LLC).

4.1.2 Models for prediction of total condensed tannins

As explained in materials and methods different wavelength regions were evaluated in order to select the best frequencies for the model construction. The most accurate determination for tannin concentration was given by partial least square regression with the two regions of 1446,42-1511,99 cm^{-1} and 2846,6-3062,6 cm^{-1} and with first derivative preprocessing method.

With this process were recognized as well two extremes values. The inclusion or exclusion of them was changing the accuracy parameter of the predictive models, so it was considered interesting to compare the two models.

In table 2 it is possible to see the values of R^2, RMSEC, RMSECV, RMSEP and RPD of the two models:

Table 2: Statistics of the prediction model without and with the two extreme values.

Model without the extreme values	Recomended Rank	Rank	R^2	RMSE	RPD
Calibration		7	83,89	216 mg/L	2,49
Cross-validation	7	7	70,78	273 mg/L	1,86
Validation	6	7	46,58	414 mg/L	1,42
Model with the extreme values					
Calibration		7	87,5	225 mg/L	2,83
Cross-validation	7	7	75,98	294 mg/L	2,04
Validation	6	7	42,43	430 mg/L	1,32

Checking the table it is possible to understand that both the model have the same PLS factors (rank=7). The value of R^2 shows a better performance for the calibration and cross-validation in the model with the extreme values (suggesting at this first sight a better model), but looking at the statistics of validation, that gives the true performance of the model, we can comprehend that the predictive model to choose is the one without the extreme values included. In fact, no matter how the performance of the model is measured
(precision, recall, MSE, correlation), we always need to measure on the test set, not on the training set. Performance on the training only tells us that the model learns what it's supposed to learn. It is not a good indicator of performance on unseen data. The test set can be obtained using an independent sample (validation set) or holdout techniques (crossvalidation). To meaningfully compare the performance of two algorithms for a given type of data, we need to compute if a difference in performance is significant.
This hypothesis is enforced by the fact that these two extremes values are considerably far from the rest of the data values, as it is easily proved by the following PCA figure.

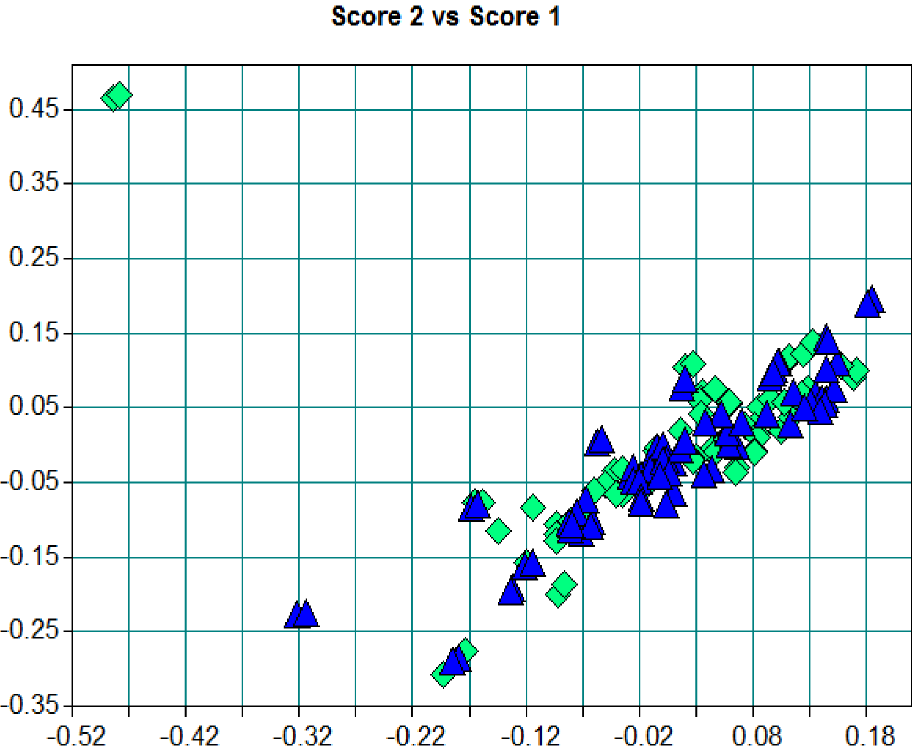

Figure 8: PCA of the samples dataset.

Chosen the model without the extreme values as the best one that can be obtained from our dataset, in reference with the accuracy and robustness of prediction, it is possible to see in the following graphs and tables if it was possible to realize a correlation between the values

determined by the reference method of the precipitation with Methylcellulose and the values predicted by this calibration model.

In this PLS regression, the dataset composed by 143 spectra (plus 8 outliers and the two extremes values) was divided in a calibration set (67 values) and a validation set (76 values) with similar standard deviations and comparable ranges as shown in table .

Table 3: Descriptive statistics of the red wine samples used for calibration and validation.

Sample	Number[a]	Range[b]	Mean[b]	SD[b]
Total	153	301-2959	1520,48	92,50
Calibration	67	496-2683	1508,37	86,57
Validation	76	301-2959	1457,95	91,79

[a] : number of samples [b] : tannin concentration in mg of CE/L. SD: Standard Deviation

We can see in graph 9, the concentration of total condensed tannins estimated by FTIR spectroscopy versus the reference values obtained by Methylcellulose method. The coefficient of correlation (R^2) was R^2=0,84, value far from 1 that cannot be included in the experiment. The value of RPD is low, RPD=2,49 and below an acceptable range, as well as the RMSEC that is relatively elevated (RMSEC=216 mg/L).

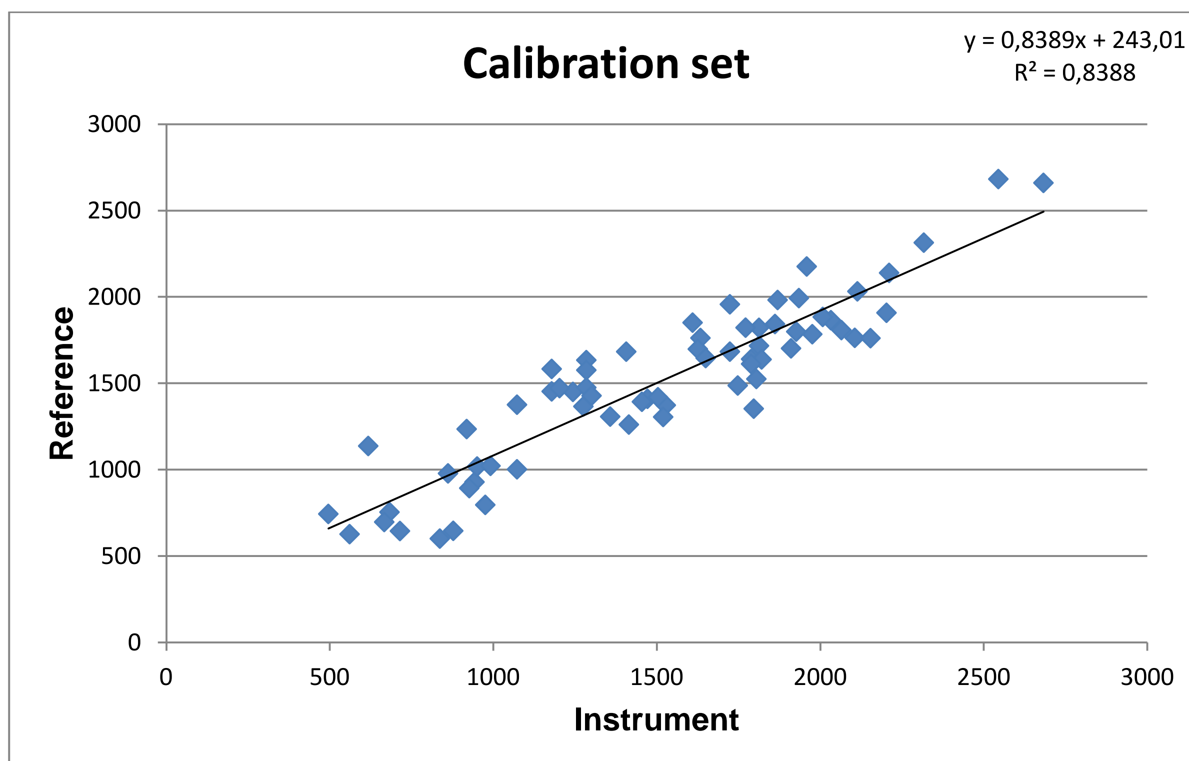

Figure 9: concentration of total condensed tannins estimated by FTIR spectroscopy instrument versus reference values from Methylcellulose method.

We can also see in figure 10, the correlation between the values of concentration of total condensed tannins predicted by the calibration model on the validation dataset versus the values determined by the method of reference on the validation dataset. The coefficient of correlation obtained was very low in this case R^2=0,51 and also the value of RPD= 1,42.

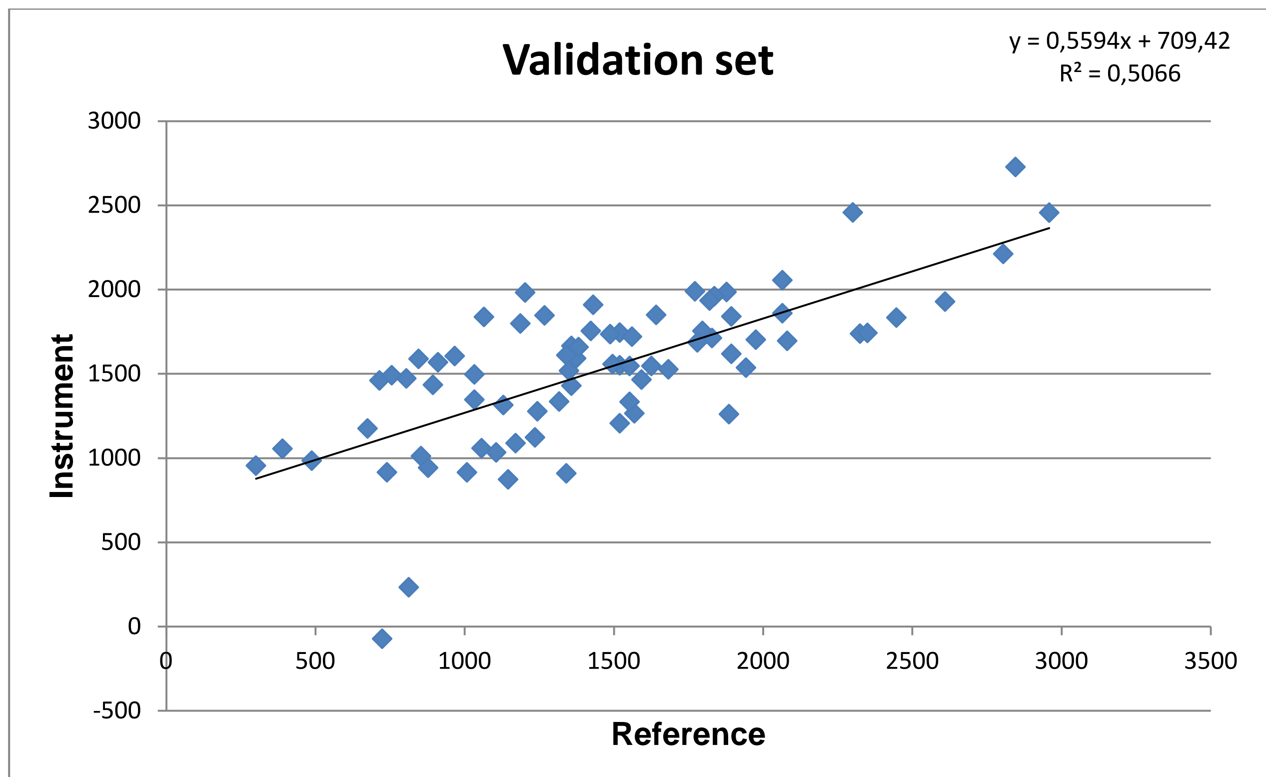

Figure 10:concentration of total condensed tannins values from the method of reference on the validation dataset versus the values predicted by the calibration model on the dataset.

The full cross-validation with 4 samples omitted, that was done, gave low values of RPD (1,86), R^2 (0,71) and a value of RMSECV of 273 mg/L.

In table 4, can be seen the statistics for total condensed tannins given by FTIR predictive model.

Table 4: statistics of the model for the prediction of condensed tannins using FT-MIR and Methylcellulose precipitation assay.

	Statistics of calibration and validation for the prediction of tannins (mg/L) using FT-MIR and Methylcellulose precipitation assay		
Samples		153	
Range		301-2959 mg/L CE	
Mean		1520,48 mg/L CE	
SD		92,50 mg/L CE	
Regions		1446,42-1511,99 cm^{-1} 2846,6-3062,6 cm^{-1}	
Rank		7	
	Calibration	Cross-validation	Validation
R^2	83,89	75,98	46,58
RPD	2,49	1,86	1,42
RMSE	216 mg/L	273 mg/L	414 mg/L

The statistics confirm that this model presents a low prediction capacity for the quantification of total condensed tannins in red wine. Especially the validation set of the model, using external values from the calibration dataset, gave a value of the coefficient of determination much lower than the ones obtained by the calibration set and cross-validation set.

The difficulties encountered in providing a good prediction model were already encountered by some authors in other studies, and are numerous and differentiated.

In fact FTIR spectroscopy presents two main limitations regarding wine analysis: similar IR absorption bands of the most interesting compounds and dominating absorption of major wine components, ethanol and water in particular. These limitation are particularly critical for the analysis of condensed tannins, because ethanol, water and organic acids are masking the IR vibrations of these phenolic compounds. (Patz *et al.,*2004; Moreira and Santos, 2004). This problem has been tried to overcome in two ways: by performing a sample pre-purification step or by using chemometrics techniques in order to take advantage of spectral regions in which there's no interference with other compounds.

Fernandez and Agosin in 2007, obtained better results in comparison with the method in this study, but Fernandez method requires an extensive sample purification step, it was performed on a single grape cultivar and the range of tannin concentration was smaller than this study (range around 85-900 mg/L CE, mean around 400 mg/L CE). In addition to this,

the method of reference with better prediction was not protein precipitation but Phloroglucinolysis. Comparing as well the results of Cozzolino et *al.* in 2008, we can see slightly better results with a R^2 of cross-validation around 0,86 and a SECV of 0,46 mg/g Epic.E and a SEP of 0,89 mg/g Epic.E; but this study was developed on homogenized red grape berries not in wine, so it is not comparable and present the disadvantage of the difficulty to predict wine tannins from grape tannin measurement due to extraction changes.

The prediction ability of the reported values by Versari *et al.* (2006), who used FTMIR Spectroscopy on direct wine solution, was very high (RMSECV = 63 mg of CE/L; *r* =0.99) but their method was developed using a high number of latent variables (10) for only 20 wines and without any independent validation of the model, the data were most likely highly overfitted.

The prediction ability of the model of Jensen *et al.* (2008) on wines showed better results but still was not really satisfactory (RMSEP= 75 mg/L CE, r= 0,94). In fact the range of tannin concentration was quite small (92-1060 mg/L CE, mean 456 mg/L CE), so if we see the results in percentage of the mean we have RMSEP(%) around 17,5% and RMSECV(%) around 13,8%, both values are higher than 10% presenting a not accurate prediction. Moreover the RPD values in calibration and in prediction were respectively 2,8 and 2,7, these RPD values would still only classify as a relatively poor prediction and would be recommended to be used just for rough screening purposes. Jensen found better quantification of tannins in grapes and considering only young red wines. In fact tannins were evaluated very well with RPD values around 5, the only problem was the low number of sample taken, just 14 sample for the validation.

Nascimento in 2011 explored the quantification of total condensed tannins by FTIR in red wines, in a range comparable with this study. In this case the model had an elevated coefficient of determination in calibration (r^2= 0,98) but the values of SECV was quite high pointing out a low predictive power of the model. Also the coefficient of determination for the validation and the SEP gave values that showed the low accuracy of the method, considering the fact, as well, that the validation was made with a low number of samples.

4.2 Single laboratory validation of the fractionation method developed by Sun *et al.*

First step for the validation process was to analyze using the same operator, the same apparatus, the same laboratory but in different days the same sample of red wine. It was chosen one sample with an average concentration of total condensed tannins in order to avoid extreme cases of this value. In Table 5, we can see the value of the three fraction (FIII: polymeric condensed tannins, FII: oligomeric condensed tannins, FI: 3-flavanoids monomeric) expressed in mg/L of epicatechin equivalents for the ten trials.

Table 5: Concentration of the three fraction of tannins for the reproducibility (mg/L Epic.E.).

	1	2	3	4	5	6	7	8	9	10
FIII	964	1045	1099,1	1027	936,9	945,9	1027	1054,1	991	973
FII	88,8	74,3	103,3	108,7	79,7	85,1	87,0	90,6	96,0	88,8
FI	8,4	9,6	9,5	8,4	9,5	7,4	10,9	8,1	8,5	10,1

In table 6, we can see the statistical analysis of the trials and the relative reproducibility in each fraction:

Table 6: statistical analysis of the reproducibility test.

	Average[a]	SD[ab]	% C.V.[c]	Reproducibility[a]	Confidence Interval[a]	Results expression[a]
FIII	1006,3	52,5	5,2	168,1	37,6	1006,3 ± 37,6
FII	90,2	10,3	11,4	32,9	7,4	90,2 ± 7,4
FI	9,0	1,0	11,6	3,4	0,8	9,0 ± 0,8

[a]: expressed in mg/L of Epicatechin Equivalents; [b]: Standard Deviation; [c]: coefficient of variance in percentage.

These values showed a good performance in line with the % of coefficient of variance obtained by Sun for the statistics of the method (% coefficient of variance FIII: 6,5; FII:6,1 and FI: 7,9). As well the confidence interval and the reproducibility can be considered good for performing these method in this laboratory. However, we don't have a target or official R to compare with our results, so that the values reported above are a mere indication referred to the experimental condition in our laboratory.

Spike recovery

Regarding the recovery rates of condensed tannins for this method, it was added in the red wine sample 5 mg/L of catechin, in order to increase the fraction FI. The sample was fractionated with the cartridge and then the three fraction of condensed tannins were quantified. After some trials where there was no instrumental perception of the spike in the quantified sample, the amount of catechin in red wine was increased to 10 mg/L (the double of the FI fraction present without spike). Also this amount was not found in the sample. So, in order to understand the reason of the lack in recovery rates, it was made a trial with a solution just with methanol and catechin at 10 mg/l and the quantity was measured directly with the vanillin. This in order to avoid the eventual loss of substance during the fractionation procedure. In this last case the quantification of the concentration of catechin in the pure methanol solution gave very low values, around 1 mg/L. This pointed out the fact that the problem of recovery was not due to the fractionation methodology process but it was a problem of the catechin reagent used.

This problem with the reagent has not permitted to explore the spike recovery for this methodology in the study.

5. CONCLUSION

This work proved the difficulties, found also in previous studies, in developing a robust model for the quantification of condensed tannins by FTIR spectroscopy directly on red wine samples. The advantages of this method over the others techniques in terms of fastness and convenience are undoubted, but the limitation mostly due to interferences in absorption of these phenolic compounds are not really overcome, even with the use of modern chemometrics techniques. In this case it was not possible to build models able to predict the concentration of condensed tannins with a significant accuracy.

To improve the accuracy of the method in the future a selection of the dataset is suggested. A wide dataset, like in this study, can give a wide range of spectra but it gives more difficulties in finding models with high fits and predictive efficacies. Dividing the dataset in samples of similar years or similar content of condensed tannins could give more performant models. In particular good results should be found using only young wines or within the same variety. Studying a fast pretreatment operation before the acquisition of the spectra is an option that could give interesting results, the elimination of some compounds in fact could give spectra with less interferences. Another option, seen the recent advances with UV/VIS Spectroscopy, is to combine it with the FT-MIR Spectroscopy and check the accuracy of the two techniques together. It is highly suggested also a study of the reference methylcellulose method in order to have a better understanding of its mechanism of action and a subsequent comparison with other methods for the quantification of condensed tannins, to decide the best methodologies to utilize as a reference .

The single laboratory validation for the fractionation method in reverse phase presented good value in percentage of coefficient of variance and reproducibility for each of the three fractions evaluated. This proved the validity and reliability of the method. The only drawback is the high amount of time and equipment required for a single sample evaluation, that could be a point where future improvements are going to be required.

6. REFERENCES

Antoine M.L., Simon C. and Pizzi A. (2004). UV spectrophotometric method for polyphenolic tannin analysis. *Journal of Applied Polymer Science 91*, 2729–2732.

Arribas M. Victoria Moreno and Polo M. Carmen (2009). Wine chemistry and biochemistry. Springer.

Bate-Smith E.C. (1973). Haemanalysis of tannins: The concept of relative astringency. *Phytochemistry 12*, 907–912.

Boulet J.C., Williams P. and Doco T. (2007). A Fourier transform infrared spectroscopy study of wine polysaccharides. *Carbohydrates Polymers, 69,* 79-85.

Broadhurst R. B., Jones W. T. (1978). Analysis of condensed tannins using acidified vanillin. *J. Sci. Food Agric.*, *29*, 788-794.

Brooks L., McCloskey L., McKesson D. and Sylvan M. (2008). Adams-Harbertson protein precipitation wine tannin method found invalid. *Journal of AOAC International, 91,* 1090-1094.

Cheynier V. and Fulcrand H. (2003). Analysis of polymeric proanthocyanidins and complex polyphenols. In „Methods of Polyphenol Analysis". Santos-Buelga, C. and Williamson, G. Eds. The Royal Society of Chemistry, Cambridge, UK, 2003. Chapter 13.

Cheynier V., Dueñas-Paton Salas E., Maury C., Souquet. J-M., Sarni-Manchado P. and Fulcrand H. (2006). Structure and properties of wine pigments and tannins. *Americal Journal of Enology and Viticulture, 57,* 298-305.

Coimbra A. C., Barros A.S., Coelho E., Gonçalves F., Rocha M.S. and Delgadillo I. (2005). Quantification of polymeric mannose in wine extract by FTIR Spectroscopy and OSC-PLS1 regression. *Carbohydrate Polymers, 61*, 434-440.

Cozzolino D., Kwiatkowski M.J., Parker M., Cynkar W.U., Dambergs R.G., Gishen M., Herderich M.J. (2004) Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta 513*, 73–80.

Da Silva Jorge M. Ricardo (1995). Estrutura e composiçao das procianidinas da uva e do vinho. Efeitos potenciais na saude*. Actas do 3° Simposio de Vitivinicultura do Alentejo*, Evora, Vol. 2.

D'Angelo R., C. Novi, F. Izzia, E. Russo, F. Cifelli. Utilizzo di un metodo chemiometrico in spettroscopia FT-IR per il dosaggio del quarzo in materie prime presenti in siderurgia. *Terzo seminario aggiornamento dei professionisti CONTARP*.

Davies and Fearn (2006). Back to basics: calibration statistics. *Spectroscopy Europe. Vol. 18 Num. 2*. 31-32.

De Beer D., Harbertson J.F., Kilmartin P.A., Roginsky V., Barsukova T., Adams D.O. and Waterhouse A.L. (2004). Phenolics: A comparison of diverse analytical methods. *American Journal of Enology and Viticulture. 55,* 389-400.

De Pascual-Teresa S., Treutter D., Rivas-Gonzalo J.C. and Santos-Buelga C. (1998). Analysis of flavanols in beverages by high-performance liquid chromatography with chemical reaction detection. *Journal of Agricultural and Food Chemistry 46,* 4209–4213.

Dixit V., Tewari J.C., Cho B. and Irudayaraj J.M. (2005). Identification and quantification of industrial grade glycerol adulteration in red wine with FTIR spectroscopy using chemometrics and artificial neural networks. *Applied Spectroscopy, 59,* 1553-1561.

Dubernet M. (2002). *Rev. Fr. Oenol.*, *181*, 10–13.

Dubernet M., Dubernet M., Grasset F. and Garcia A. (2001). Analyse de l'azote assimilable dans le mouts par Interferometrie Infrarouge a Transformee de Fourier. Revue Française d'œnologie, 187, 9-13.

Edelmann A., Diewok J., Schuster K.C. and Lendl B. (2001). Rapid method for the discrimination of red wine cultivars based on mid-infrared spectroscopy of phenolic wine extracts. *J. Agric. Food Chem., 49,* 1139-1145.

Esbensen *et al.,* (2008) FTIR spectroscopy for grape and wine analysis. *Analytical Chemistry, March, 1.* 1372-1379.

Fernandez K. and Agosin E. (2007). Quantitative Analysis of Red Wine Tannins Using Fourier-Transform Mid-Infrared Spectrometry. *Journal of Agric. Food Chem., 55,* 7294-7300.

Fu Y., Lim L.T. and McNicholas (2009). Changes on enological parameters of white wine Packaged in Bag-in-box during secondary shelf life. *Journal of food science, 74*, n°8.

Gallignani M., Ayala C., Brunetto M.R., Burguera J.L. and Burguera M. (2005). A simple strategy for determining ethanol in all types of alcoholic beverages based on its on-line liquid-liquid extraction with chloroform, using a flow injection system and Fourier transform infrared spectrometric detection in the mid-IR. *Talanta, 68*, 470-479.

Guyot S., Doco T., Souquet J.M., Moutounet M., Drilleu J.F., (1997). Characterization of highly polymerized procyanidins in cider apple (Malus Sylvestris Var Kermerrien) skin and pulp. *Phytochemistry, 44 (2),* 351-357.

Hai-bin Qu, Dan-lin Ou and Yi-yu Cheng (2005). Background correction in near-infrared spectra of plant extracts by orthogonal signal correction. *J. Zhejiang Univ. Sci. B. August; 6(8):* 838-843.

Hayasaka Y., Waters E.J., Cheynier V., Herderich M.J., Vidal S. (2003). Characterization of proanthocyanidins in grape seeds using electrospray mass spectrometry. *Rapid communications in mass spectrometry, 17,* 9-16.

Harbertson J.F., Picciotto E.A. and Adams D.O. (2003). Measurement of polymeric pigments in grape berry extracts and wines using a protein precipitation assay combined with bisulfite bleaching. *Americal Journal of Enology and Viticulture, 54,* 301-306.

Herderich M.J. and Smith P.A. (2005). Analysis of grape and wine tannins: Methods, applications and challenges. *Australian Journal of Grape and Wine Research, 11,* 205–214.

Jackson Ronald S. (2008). Wine Science: principles and applications. Third Edition. Elsevier Inc.

Jensen Jacob S., Max Egebo and Anne S. Meyer (2008). Identification of Spectral Regions for the Quantification of Red Wine Tannins with Fourier Transform Mid-Infrared Spectroscopy. *J. Agric. Food Chem.*, *56,* 3493–3499.

Jones W.T. and Mangan J.L. (1977). Complexes of the condensed tannins of sainfoin (*Onobrychis viciaefolia Scop*) with fraction 1 leaf protein and with submaxillary mucoprotein and their reversal by PEG and pH. *Journal of the Science of Food and Agriculture 28,* 126–136.

Kennedy J. A. (2008) Grape and wine phenolics: Observations and recent findings. *Cien. Inv. Agr. 35(2)*:107-120.

Lee C.Y., Jaworski A.W. (1987). Phenolic compounds in white grapes grown in New York. *Am. J. Enol. Vitic., 38,* 277-281.

Le Roux E., Doco T., Sarni P., Lozano Y., Cheynier V. (1998). A-type proanthocyanidins from pericarp of Litchi Chinensis. *Phytochemistry, 48,* 1251-1258.

Luck G., Liao H., Murray N. J., Grimmer H. R., Warminski E. E., Willamson M. P., Lilley T.H., & Haslam E. (1994). Polyphenols, astringency and prolin-rich proteins. *Phytochemistry,37*, 357–371.

Makkar H.P.S., Blummel M., Borowy N.K. and Becker K. (1993). Gravimetric-determination of tannins and their correlations with chemical and protein precipitation methods. *Journal of the Science of Food and Agriculture 61,* 161–165.

Martens H. and Naes T. (1989). Multivariate calibration. John Wiley and Sons Ed.

McMurrough I. and McDowell J. (1978). Chromatographic separation and automated analysis of flavanols*. Analytical Biochemistry 91,* 92–100.

Mercurio M.D., Dambergs R.G., Herderich M.J. and Smith P.A. (2007). High throughput analysis of red wine and grape phenolics – adaption and validation of methyl cellulose precipitable tannin assay and modified Somers color assay to a rapid 96 well plate format. *Journal of Agricultural and Food Chemistry, 55,* 4651-4657.

Mercurio M. and Smith P. (2008). Tannin quantification in red grapes and wine: Comparison of polysaccharide- and protein-based tannin precipitation techniques and their ability to model wine astringency. *Journal of Agricultural and Food Chemistry, 56,* 5528-5537.

Montedoro G. and Fantozzi P. (1974). Dosage des tannins dans les moûts et les vins a l'aide de la methylcellulose et evalution d'autres fractions phénoliques*. Lebensmittel-Wissenschaft und-Technologie 7,* 155–161.

Moreira J. L., Marcos A. M., Barros P. (2002), *Ciêcia Téc. Vitiv.*, *17*, 27–33.

Moreira J. L., Santos L. (2004). Spectroscopic interferences in Fourier transform infrared wine analysis. *Anal. Chim. Acta*, *513* (1), 263–268.

Moreira J.L. and Santos L. (2005). Analysis of organic acids in wines by Fourier-transform infrared spectroscopy. *Anal. Bioanal. Chem.* 382, 421-425.

Nascimento A. P. R. (2011). Quantificaçao dos tanninos dos vinhos por transformada de Fourier dos espectros no infravermelho médio (FTIR). Master dissertation. Instituto Tecnico de Lisboa-UTL, Lisboa

Oszmianski J., Ramos T., Bourzeix M. (1988). Fractionation of phenolic compounds in red wine. *Am. J. Enol. Vitic., 39 (3),* 259-262 .

Ozgur Yeniay and Atilla Goktas (2002). A comparison with partial least square regression with other predictive methods. *Hacettepe Journal of Mathematics and Statistics Volume 31*, 99-111.

Passos Cláudia P. *et al.* (2010). Application of Fourier transform infrared spectroscopy and orthogonal projections to latent structures/partial least squares regression for estimation of procyanidins average degree of polymerization. *Analytica Chimica Acta 661,* 143–14.

Patz C.D. *et al.* (2000). Wine analysis using FTIR. *In focus 24, (1),* 16-18.

Patz C. D., Blieke A., Ristow R., Dietrich H. (2004). Application of FT-MIR spectrometry in wine analysis. *Anal. Chim. Acta*, *513* (1), 81–89.

Peyrot des Gachons C. and J.A. Kennedy (2003). Direct method for determining seed and skin proanthocyanidin extraction in red wine. *J. Agric. Food Chem. 51*:5877-5881.

Pizarro C., Gonzales Saiz J.M., Esteban Diez I. and Orio P. (2011). Prediction of total and volatile acidity in red wines by Fourier-transform mid-infrared spectroscopy and iterative predictor weighting. *Anal Bioanal Chem., 399,* 2061-2072.

Porter L.J., Hrstich L.N. and Chan B.G. (1986). The conversion of procyanidins and prodelphinidins to cyanidin and delphinidin. *Phytochemistry 25,* 223–230.

Price M. L., Scoyoc S. V., Butler L. G. (1978). A critical evaluation of the vanillin reaction as an Assay for tannin in sorghum grain. *J. Agric. Food Chem. 26*, 1214-1218.

Rasmus Bro (2003). Multivariate calibration. What is in chemometrics for the analytical chemist? *Analytica Chimica Acta 500,* 185–194

Reed J.D., Horvath P.J., Allen M.S. and Vansoest P.J. (1985). Gravimetric-determination of soluble phenolics including tannins from leaves by precipitation with trivalent ytterbium. *Journal of the Science of Food and Agriculture 36,* 255–261.

Revilla E., Alonso E., Bourzeix M., Heredia N. (1991). Dosage des catéchines et des proanthocyanidols dans les vins. *Feuillet Vert O.I.V.,* 829.

Ribereau-Gayon P., Y. Glories, A. Maujean, D. Dubourdieu. (2006). Handbook of Enology Volume 2 The Chemistry of Wine Stabilization and Treatments 2nd Edition. Wiley.

Ribéreau-Gayon P., P. Sudraud, J.C. Milhé, and A. Canbas (1970). Reserches technologiques sur les composes phenoliques des vin rouges II . Les factors de dissolution des composes phénoliques. *Conn. Vigne Vin 4:*133-144.

Rigaud J., Escribano-Bailon M.T., Prieur C., Souquet J.M. and Cheynier V. (1993). Normal-phase high-performance liquid chromatographic separation of procyanidins from cacao beans and grape seeds. *Journal of Chromatography A 654,* 255–60.

Santos-Buelga C. and de Freitas V. (2009). Influence of phenolics on wine organoleptic properties. In „Wine Chemistry and Biochemistry", Moreno-Arribas, M.V and Polo, M.C. Eds., Springer, New York, 2009. Chapter 9D.

Sarneckis, Dambergs, Jones, Mercurio, Herderich & Smith (2006). Quantification of condensed tannins by precipitation with methyl cellulose: development and validation of an optimised tool for grape and wine analysis. *Australian Journal of Grape and Wine Research* 12, 39–49, 2006.

Saucier C., Mirabel M., Daviaud F., Longieras A., Glories Y. (2001). Rapid fractionation of grape seed proanthocyanidins. *J. Agric. Food Chem., 49,* 5732-5735.

Scalbert A. (1992). Quantitative mehods for estimation of tannins in plant tissues. In R. W. Hemingway and P.S.Laks, Plant Polyphenols: Synthesis, properties, significance. Plenim Press.

Schneider V. (1995). Evaluation of small amounts of flavonoid phenols in white wines by colorimetric assays. *American Journal of Enology and Viticulture 46,* 274–277.

Schneider R., Charrier F., Moutounet M. and Baumes R. (2004) Rapid analysis of grape aroma glycoconjugates using FTIR spectrometry and chemometric techniques. *Anal. Chim. Acta, 513,* 91-96.

Scollary Geoffrey R. (2010) .Tannin Review. Grape and wine research & development corporation.

Seddon T.J. and Downey M.O. (2008). Comparison of analytical methods for the determination of condensed tannins in grape skin. *Australian Journal of Grape and Wine Research* 14, 54–61, 2008

X J. Shang *et al.* Jinling Hospital, Nanjing University Clinical School of Medicine, Nanjing, P. R. China. (2008). Procyanidin Induces Apoptosis and Necrosis of Prostate Cancer Cell Line PC-3 in a Mitochondrion-Dependent Manner.

Sherman C.P. Hsu, Ph.D. Handbook of Instrumental Techniques for Analytical Chemistry.

Shiuan Chen, Ph.D. Beckman Research Institute of the City of Hope, Duarte California. ES8258.

Silanikove N., Shinder D., Gilboa N., Eyal M. and Nitsan Z. (1996). Binding of polyethylene glycol to samples of forage plants as an assay of tannins and their negative effects on rumenal degradation. *Journal of Agricultural and Food Chemistry 44,* 3230–3234.

Sinelli N. Introduzione all'analisi multivariata dei dati. Dipartimento Scienze e Tecnologie Alimentari e Microbiologiche. Sezione Tecnologie Alimentari. Università degli Studi di Milano.

Skoutelas D. (2010). Assimilable nitrose of wine grapes. Validations and comparison of formol and FTIR method. Master dissertation. Instituto Tecnico de Lisboa-UTL, Lisboa

Soriano A., Perez Juan P., Vicario A., Gonzalez J. And Perez Coello M. (2007). Determination of anthocyanins in red wine using a newly developed method based on FTIR spectroscopy. *Food Chemistry 104,* 1295-1303.

Stevanato R., Fabris S. and Momo F. (2004). Enzymatic method for the determination of total phenolic content in tea and wine. *Journal of Agricultural and Food Chemistry 52,* 6287–6293.

Sun Baoshan, Jorge M. Ricardo-da-Silva and Isabel Spranger (1998b). Critical Factors of Vanillin Assay for Catechins and Proanthocyanidins. *J. Agric. Food Chem. 46,* 4267-4274.

Sun B.S., Leandro M.C., Ricardo-da-Silva J.M., Spranger M.I., (1998a). Separation of grape and wine proanthocyanidins according to their degree of polymerization. *J. Agric. Food Chem. 46,* 1390-1396.

Sun Baoshan, M. Isabel Spranger (2005). Review: quantitative extraction and analysis of grape and wine proanthocyanidins and stilbenes. *Ciência Téc. Vitiv. 20 (2),* 59-89. 2005.

Swanepoel M., du Toit M. and Nieuwoudt H.H. (2007). Optimisation of the quantification of total soluble solids, pH and titratable acidity in South African grape must using FTIR Mid Infrared spectroscopy. *S. Afr. J. Enol. Vitic., 28 (2),* 140-149.

Tauler R., E. Peré-Trepat, S. Lacorte, D. Barcelò. Chemometrics Modelling of Environmental Data. *Department of Environmental Chemistry, Institute of Chemical and Environmental Research,* IIQAB-CSIC.

Treutter D. (1989). Chemical-reaction detection of catechins and proanthocyanidins with 4 dimethylaminocinnamaldehyde. *Journal of Chromatography 467,* 185–193.

Treutter D., Santos-Buelga C., Gutmann M. and Kolodziej H. (1994). Identification of flavan-3-ols and procyanidins by high-performance liquid-chromatography and chemical-reaction detection. *Journal of Chromatography A 667,* 290–297.

Urtubia A., Perez Correa J.R., Meurens M. and Agosin E. (2004). Monitoring large scale wine fermentations with infrared spectroscopy. *Talanta, 64,* 778-784.

Versari, A.; Boulton, R. B.; Parpinello, G. P. (2006). Effect of spectral pre-processing methods on the evaluation of the color components of red wines using Fourier-transform infrared spectrometry. *Ital. J. Food Sci., 18* (4), 423–431.

Versari A., Laghi L., Thorngate J.H. and Boulton R.B. (2011). Prediction of colloidal stability in white wines using infrared spectroscopy. *Journal of food engineering, 104*, 239-245.

Versari a., Parpinello G.P., Scazzina F. and Del Rio D. (2010). Prediction of total antioxidant capacity of red wine by Fourier transform infrared spectroscopy. *Food control, 21*, 786-789.

Yanagida A., Shoji T. and Shibusawa Y. (2003). Separation of proanthocyanidins by degree of polymerization by means of size-exclusion chromatography and related techniques. *Journal of Biochemical and Biophysical Methods 56,* 311–322.

Waters W.A. (1964). Mechanisms of Oxidation of Organic Compounds, Vol. 6, John Wiley and Sons Inc., New York, pp. 28–138.

Williams P.C. and D.C. Sobering. (1993). Comparison of commercial Near Infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *J. Near infrared spectrosc.1,* 25-32.

Wolfender J-L., Ndjoko K. and Hostettmann K. (2003). Application of LC-NMR in the structure elucidation of polyphenols. In „Methods of Polyphenol Analysis". Santos-Buelga, C. and Williamson, G. Eds. The Royal Society of Chemistry, Cambridge, UK, 2003. Chapter 6.

Wynne L., Clark S., Adams M. and Barnett N. (2007). Compositional dynamics of a commercial wine fermentation using two-dimensional FTIR correlation analysis. *Vibrational spectroscopy, 44,* 394-400.

Web sites:

http://www.ansci.cornell.edu/plants/toxicagents/tannin.html
www.en.wikipedia.org.
http://mmrc.caltech.edu/FTIR/FTIRintro.pdf
www.chemiometria.it

www.Wiki.eigenvector.com

Annexes: value of concentration of total condensed tannins in red wines expressed in mg/L of catechin equivalents.

Number	Sample	Total condensed tannins
1	Cuba Sial 2011	1358
2		1374
3	Castelao Isa 2010	715
4		805
5	Periquita de Fonseca 1969	1789
6		1813
7	Casa Agricola Fernando Mourc Dep 102 Alicante Testimunha	2268
8		1325
9	Casa agricola Fernando Mourc Dep 102 Alicante Enscio 5% conc	1772
10		1878
11	Casa agricola Fernando Mourc Dep 53 Enscio 15% conc	1285
12		1797
13	Real comp velha Cuba f62 TN 15% conc	1317
14		1837
15	Pinot Noir 2011 Barrica 4	1146
16		1341
17	Casa agricola Fernando Mourc Dep 53 Enscio Testimunia	1033
18		967
19	Merlot 2009 Barrica 4	618
20		1073
21	Merlot 2009 Barrica 2	1171
22		1244
23	Casa agricola Fernando Mourc Dep 102 Alicante Enscio 10% conc	2106
24		1772
25	Tinto 1 2011	1130
26		1236
27	Tinto 2 011	943
28		976
29	Casa agricola Fernando Mourc Dep 53 Enscio 10% conc	1472
30		1407
31	Real comp velha Cuba f62 TN 20% conc	2033
32		2211
33	Casa agricola Fernando Mourc Dep 71 Alicante Enscio 10% conc	1894
34		1821
35	Casa agricola Fernando Mourc Dep 102 Alicante Enscio 15% conc	2325
36		2081
37	Alentejo 2010 Alicante Alfrocheiro Aragones Herdade do Pico	1976
38		1935
39	Casa agricola Fernando Mourc Dep 71 Alicante Enscio 15% conc	1813
40		2154
41	Merlot 2009 Barrica 1	951
42		919
43	Quinta da Silveira Douro 2011	846
44		1203

45	Annuda 2011 Cuba A1	1911
46		1748
47	Annuda 2011 Cuba 500L	1943
48		1886
49	Annuda 2011 Cuba nova	1179
50		1285
51	Annuda 2011 Cuba C1	2114
52		1959
53	Annuda 2011	1341
54		1520
55	Annuda 2011 Cuba C2	1821
56		1927
57	Syrah Cabernet 2008	894
58		756
59	Tinto 3	528
60		642
61	Castelao 2011	732
62		821
63	Syrah Touriga 2011	1358
64		1382
65	Annuda 2011 Cuba A2	1829
66		1797
67	Syrah Touriga Cabernet 2011	1187
68		1065
69	Tinto 2	488
70		675
71	Syrah Touriga Cabernet 2011 Den	1610
72		1724
73	Tinto 1	1520
74		1626
75	Touriga 2011	1244
76		1203
77	Trincadeira 2011	1285
78		1276
79	Trincadeira Castelao 2011	984
80		976
81	Cabernet 2011	1000
82		740
83	Syrah 2011	837
84		862
85	Merlot 2009 Barrica 3	1033
86		911
87	Real comp velha Cuba f62 TN amostra original	2203
88		2008
89	Real comp velha Cuba f62 TN 5% conc	2447
90		2350
91	Casa agricola Fernando Mourc Dep 53 Enscio 5% conc	1431
92		1268
93	M. Annochais M LO T 5 M	2317
94		2333
95	Tinto	3407
96		3577

97	M. Annochais 4/2	2545
98		2683
99	M. Annochais M T 7,5 M	2846
100		2301
101	M. Annochais 3/2	2846
102		2959
103	M. Annochais M T 5M	2065
104		2065
105	Touriga nacional ISA 2011	813
106		724
107	Aragones N14 Canmo	1976
108		1894
109	Pegoes F1 2002	854
110		740
111	Pegoes NR 2000	683
112		715
113	Aragones Evonc R2P3	2610
114		2805
115	Trincadeira ISA 2000	301
116		390
117	Touriga Nacional CFT0 I 2010	1520
118		1317
119	Estremoz Aragonez 2	1870
120		2065
121	Pinot Noir 2010 CFT 03	561
122		496
123	Pinot Noir 2010	878
124		667
125	Touriga Nacional 2011	2390
126		2415
127	Pinot Noir B1 2011	927
128		992
129	Merlot 2011	1488
130		1683
131	Pinot Noir Bq 2011	1106
132		1008
133	Atestos	1789
134		1862
135	Ramiscos 2011	1268
136		1366
137	Touriga nacional CFT 01	1073
138		1358
139	Monte de greis 2007 Aragones DI MOB	602
140		642
141	Monte de greis 2007 Aragones N1	1179
142		1187
143	Aragones CF5 2003 Marco	1455
144		1528
145	Touriga nacional Liz ISA 2003	1593
146		1569
147	Syrah Liz ISA 2003	878
148		1057

149	Aragones CN 11 2003 Marco	1520
150		1504
151	Prof Arruda 8 2008	1512
152		1415
153	Paula F Syrah Minho 2008	390
154		390
155	Paula F Syrah Cabecao 2008	797
156		667
157	Paula F Touriga Vidiguena 2008	1098
158		1293
159	Vinho 1 Ines O	1634
160		1626
161	Paula F. Touriga Leira 2008	1350
162		1553
163	Vinho 2 Ines O	1496
164		1553
165	Ava Rita Di Mob 2008	1301
166		1415
167	Ava Rita Di Rel 2008	1805
168		1724
169	Tinto 3 9/4/2012	1642
170		1423
171	Tinto	1780
172		1561
173	Ana Rita Pro Rel 2008	1650
174		1610
175	Prof Aruda 2007 3	545
176		504