



**HAL**  
open science

# Appréciation des tanins, de la couleur et de l'astringence des raisins, moûts et vins par technologies analytiques IRTF et UV-visible couplées à l'analyse de régression multivariée

Clément Miramont

## ► To cite this version:

Clément Miramont. Appréciation des tanins, de la couleur et de l'astringence des raisins, moûts et vins par technologies analytiques IRTF et UV-visible couplées à l'analyse de régression multivariée. Chimie analytique. Université de Bordeaux, 2021. Français. NNT : 2021BORD0091 . tel-03216487

**HAL Id: tel-03216487**

**<https://theses.hal.science/tel-03216487v1>**

Submitted on 4 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE PRÉSENTÉE  
POUR OBTENIR LE GRADE DE  
**DOCTEUR DE**  
**L'UNIVERSITÉ DE BORDEAUX**

ÉCOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ  
SPÉCIALITÉ ŒNOLOGIE

Par Clément MIRAMONT

**Appréciation des tanins, de la couleur et de l'astringence des  
raisins, moûts et vins par technologies analytiques IRTF et UV-  
visible couplées à l'analyse de régression multivariée**

Sous la direction du Professeur Pierre-Louis TEISSEDRE

Soutenue le 24 mars 2021

Membres du jury :

Pr. ZAMORA, Fernando  
Pr. KALLITHRAKA, Stamatina  
Pr. DE FREITAS, Victor  
Pr. MARÍN ARROYO, María Remedios  
Pr. TEISSEDRE, Pierre-Louis  
M. THIIIS HEIDE, Søren

Université Rovira i Virgili  
Université d'Athènes  
Université de Porto  
Université Publique de Navarre  
Université de Bordeaux  
Foss Analytical A/S

Président  
Rapporteuse  
Rapporteur  
Examinatrice  
Directeur de thèse  
Membre invité



## **Appréciation des tanins, de la couleur et de l'astringence des raisins, moûts et vins par technologies analytiques IRTF et UV-visible couplées à l'analyse de régression multivariée**

L'appréciation analytique rapide (quantitative et qualitative) des composés phénoliques des raisins, des moûts et des vins rouges et leur incidence organoleptique (astringence, amertume) apparaît être un challenge en œnologie. Plusieurs méthodes analytiques ont été développées dans cette optique, mais elles nécessitent du matériel, du temps et des connaissances qui ne permettent pas de les appliquer efficacement en routine dans le milieu industriel. Pour optimiser la production des vins rouges en lien avec les concentrations et compositions en anthocyanes et en tanins, nous avons développé des modèles de prédiction fiables grâce à l'IRTF et l'UV-visible couplés à l'analyse de régression multivariée PLS.

Afin de répondre à ces attentes, un ensemble d'échantillons représentatifs de la maturité du raisin, de la fermentation alcoolique, de l'élevage en barriques ainsi que des vins commerciaux a été collecté. Ces échantillons ont été analysés grâce à différentes méthodes décrites dans la littérature, par réaction chimique ou par analyse CLHP-UV-visible, afin d'obtenir la composition et la concentration en flavan-3-ols, la concentration et la composition en anthocyanes, la composition en pigments monomériques et polymériques ainsi que la concentration en tanins totaux, précipitables à l'albumine de sérum bovin et la méthylcellulose. Une sélection de vins a également été testée par un panel de dégustateurs entraînés afin de déterminer un indice d'astringence et d'amertume, avec enregistrement des spectres IRTF et UV-visible de ces échantillons.

La corrélation entre les analyses spectrales et les différentes informations analytiques obtenues a été recherchée grâce à l'analyse de régression multivariée PLS, dans le but de concevoir des modèles de prédiction. Les différentes modélisations ont ensuite été testées grâce à une validation croisée, et une validation avec un jeu d'échantillons externes à la calibration.

Les résultats obtenus démontrent l'intérêt de l'IRTF et de l'UV-visible couplés à l'analyse de régression PLS pour prédire efficacement les concentrations et compositions en anthocyanes et en tanins des raisins, moûts et vins. Il apparaît que l'UV-visible est plus adapté à la modélisation des concentrations globales, alors que l'IRTF montre un fort potentiel pour une caractérisation plus approfondie de la composition moléculaire. De plus, l'addition à l'IRTF de longueurs d'onde du visible spécifiques a permis d'augmenter la robustesse des modèles pour la prédiction de l'évolution des anthocyanes. Pour la sensation d'astringence, de bons résultats préliminaires ont été obtenus pour l'IRTF et l'UV-visible. L'amertume, en revanche, s'est avéré être une perception trop complexe pour être modélisée.

Les limites de prédiction des deux méthodes spectroscopiques ont aussi pu être discutées, notamment la forte variabilité induite par les différents millésimes étudiés qui peut fortement impacter les résultats de prédiction et doit être prise en compte, ainsi que la surinterprétation de résultats prédictifs pour les concentrations en anthocyanes moléculaires qui peuvent amener à questionner certaines modélisations présentes dans la littérature.

Nos travaux permettent de mieux appréhender les avantages et les limites de l'IRTF et l'UV-visible couplés à l'analyse de régression PLS pour le dosage des anthocyanes et tanins des raisins, moûts et vins. La possibilité de modéliser grâce à une analyse spectrale une perception telle que l'astringence de façon innovante a aussi été étudiée pour la première fois. L'ensemble des résultats obtenus ont permis d'approfondir les connaissances actuelles sur le sujet et de développer de nouveaux modèles de prédiction fiables, efficaces et transférables pour un usage industriel dans le domaine de la vigne et du vin.

**Mots clés :** Polyphénols, Infrarouge à transformée de Fourier, UV-visible, Analyse de régression PLS, Tanins, Anthocyanes

---

**Unité de recherche Œnologie**

EA 4577, USC 1366 INRAE, IPB - 210 Chemin de Leysotte, CS 50008

33882 Villenave d'Ornon Cedex

## **Tannins, color and astringency appreciation for grapes, musts and wines by quick analytical technologies FTIR and UV-visible with multivariate regression analysis**

The rapid analytical appreciation (quantitative and qualitative) of the phenolic compounds of grapes, musts and red wines and their organoleptic incidence (astringency, bitterness) appears to be a challenge in oenology. Several analytical methods have been developed, but they require equipment, time and knowledge which do not allow them to be applied effectively in routine for industrial environment. To optimize the production of red wines in relation to anthocyanins and tannins concentrations and compositions, we have developed reliable prediction models for FTIR and UV-visible with PLS multivariate regression analysis.

In order to obtain these expectations, a set of representative samples from grapes maturity, alcoholic fermentation, aging in barrels and commercial wines were collected. These samples were analyzed using different methods described in the literature, by chemical reaction or by HPLC-UV-visible analysis, in order to obtain the composition and concentration of flavan-3-ols, concentration and composition of anthocyanins, composition of monomeric and polymeric pigments and concentration of total tannins, bovine serum albumin and methylcellulose precipitable tannins. A selection of wines was also tested by a panel of trained tasters to determine an astringency and bitterness index, with FTIR and UV-visible spectra recording of these samples.

Correlation between spectral analyzes and the various analytical information obtained were sought with PLS multivariate regression analysis, for designing prediction models. The different models were tested with cross validation, and validation with an external set of samples to the calibration.

The results obtained demonstrate the interest of FTIR and UV-visible with PLS multivariate regression analysis to effectively predict anthocyanins and tannins concentrations and compositions in grapes, musts and wines. It appears that UV-visible is more suited to modeling global concentrations, while FTIR shows great potential for the characterization of molecular composition. The addition of specific visible wavelengths to FTIR has increased model robustness for predicting anthocyanin evolution. For astringency sensation, good preliminary results were obtained for FTIR and UV-visible. Bitterness turned out to be a too complex perception to be modeled.

Prediction limits of the two spectroscopic methods were also discussed, especially the strong variability induced by different vintages studied who can strongly impact the prediction results and should not be overlooked, as well as the overinterpretation of predictive results for molecular anthocyanin concentrations may lead to questioning some models in the literature.

Our work provides better understanding of the FTIR and UV-visible advantages and limits with PLS multivariate regression analysis for determination of anthocyanins and tannins in grapes, musts and wines. The possibility of modeling a perception such as astringency in an innovative way using spectral analysis was also investigated for the first time. All the results obtained have deepened the current knowledge on the subject and they allowed to develop new reliable and efficient prediction models, transferable for industrial use in the vine and wine sector.

**Keywords** : Polyphenols, Fourier transform infrared, UV-visible, Partial least squares regression, Tannins, Anthocyanins

---

**Unité de recherche Œnologie**

EA 4577, USC 1366 INRAE, IPB - 210 Chemin de Leysotte, CS 50008

33882 Villenave d'Ornon Cedex

## Valorisation des travaux de recherches

### Publications internationales :

Miramont, C., Jourdes, M., Selberg, T., Juhl, H. V., Nørgaard, L., & Teissedre, P. L. (2019). Prediction of anthocyanin concentrations during red wine fermentation: development of Fourier transform infrared spectroscopy partial least squares models. *OENO One*, 53(4)

Miramont, C., Jourdes, M., & Teissedre, P. L. (2020). Development of UV-vis and FTIR Partial Least Squares models: comparison and combination of two spectroscopy techniques with chemometrics for polyphenols quantification in red wine. *OENO One*, 54(4), 779-792.

González-Centeno, M. R., Chira, K., Miramont, C., Escudier, J. L., Samson, A., Salmon, J. M., ... & Teissedre, P. L. (2019). Disease resistant bouquet vine varieties: Assessment of the phenolic, aromatic, and sensory potential of their wines. *Biomolecules*, 9(12), 793.

### Communications orales internationales :

Oeno-Ivas 2019, Bordeaux, France : Miramont, C., Jourdes, M., Selberg, T., Juhl, H. V., Nørgaard, L., & Teissedre, P. L. « Development of FTIR Partial Least Squares models for polyphenol quantification in red wine during fermentation »

OIV 2019, Genève, Suisse : Miramont, C., Jourdes, M., Selberg, T., Juhl, H. V., Nørgaard, L., & Teissedre, P. L. « Prediction of anthocyanin level during red wine fermentation: development of FTIR Partial Least Squares models »

### Communications écrites internationales :

Macrowine 2018, Saragosse, Espagne : Miramont, C., Jourdes, M., Selberg, T., Nikolajsen, T., Teissedre, P. L. « Development of UV-vis and FTIR models for polyphenol quantification in red wine »

OIV 2018, Punta del Este, Uruguay : Miramont, C., Jourdes, M., Selberg, T., Nikolajsen, T., Teissedre, P. L. « Development of UV-vis models for polyphenol quantification in red wine »

Oeno-Ivas 2019, Bordeaux, France : Miramont, C., Jourdes, M., Selberg, T., Juhl, H. V., Nørgaard, L., & Teissedre, P. L. « Development of FTIR Partial Least Squares models for polyphenol quantification in red wine during fermentation »

# Table des matières

<b>Introduction et objectifs .....</b>	<b>1</b>
<b>I. Etude bibliographique .....</b>	<b>7</b>
A. Les polyphénols du vin : composition, évolution et impact sensoriel .....	9
1. Les composés non-flavonoïdes .....	10
a) Les acides phénols .....	10
b) Les stilbènes .....	13
2. Les composés flavonoïdes .....	14
a) Les flavones, les flavanones, les flavonols et les flavononols .....	15
(1) Les flavones.....	15
(2) Les flavanones.....	15
(3) Les flavonols .....	16
(4) Les flavononols .....	17
b) Les flavan-3-ols et tanins condensés.....	18
(1) Les monomères de flavan-3-ols .....	18
(2) Les tanins condensés ou proanthocyanidines .....	19
(3) Les tanins couronnes .....	22
c) Les anthocyanes .....	23
(1) Structures, localisations et concentrations.....	23
(2) Equilibre des anthocyanes en fonction du pH .....	25
(3) Décoloration des anthocyanes en présence de bisulfite de sodium.....	26
3. Les tanins hydrolysables .....	27
a) Classification et structures des tanins hydrolysables .....	27
b) Extraction et évolution des ellagitanins dans le vin.....	29
c) Réaction des ellagitanins avec les flavonoïdes .....	30
4. Evolution des flavonoïdes du vin : condensation et dégradation.....	32

a)	Réactions directes de condensation.....	32
(1)	Polymérisation directe par catalyse acide .....	32
(2)	Polymérisation directe en présence d'oxygène .....	34
(3)	Polymérisation directe entre les flavan-3-ols et les anthocyanes .....	34
b)	Réactions indirectes de condensation en présence d'acétaldéhyde .....	35
c)	Evolution des anthocyanes dans les vins .....	37
(1)	Dégradation des anthocyanes .....	37
(2)	Copigmentation des anthocyanes .....	38
(3)	Formation de pyranoanthocyanes.....	38
d)	Impact des macromolécules du vin sur les tanins condensés.....	40
5.	Propriétés sensorielles des polyphénols du vin.....	41
a)	Perception d'astringence .....	41
b)	Perception d'amertume .....	42
c)	Astringence et amertume des polyphénols du vin .....	43
B.	La spectroscopie UV-visible et IRTF appliquée aux polyphénols du raisin, du moût et du vin	45
1.	La spectroscopie UV-visible appliquée aux polyphénols du raisin, du moût et du vin.....	46
a)	La spectroscopie UV appliquée aux tanins condensés.....	47
b)	La spectroscopie visible appliquée aux anthocyanes .....	48
2.	L'infrarouge à transformée de Fourier appliqué aux polyphénols du raisin, du moût et du vin	49
a)	L'infrarouge à transformée de Fourier appliqué aux polyphénols.....	50
C.	L'analyse de données multivariée : généralités et application aux polyphénols du raisin, du moût et du vin .....	53
1.	La régression PLS (régression par les moindres carrés partiels) .....	54
a)	Description de l'analyse et la régression PLS.....	54
b)	Prétraitement des données spectrales.....	56

c)	Etapes de construction d'un modèle de prédiction .....	56
d)	Validation d'un modèle de prédiction.....	57
2.	Paramètres descriptifs d'un modèle de prédiction .....	59
3.	Application de la spectroscopie couplée à la régression PLS pour l'analyse des polyphénols du raisin, du moût et du vin.....	60
a)	Utilisation de la spectroscopie UV-visible pour l'analyse des polyphénols du raisin, du moût et du vin .....	60
(1)	L'UV-visible combiné au proche infrarouge .....	60
(2)	L'UV-visible .....	61
b)	Utilisation de la spectroscopie infrarouge pour l'analyse des polyphénols du vin.....	63
D.	Conclusion générale.....	65
<b>II.</b>	<b>Matériels et méthodes .....</b>	<b>67</b>
A.	Matériels.....	69
1.	Réactifs, solvants et standards .....	69
2.	Echantillonnage.....	70
a)	Echantillons de raisin .....	70
b)	Echantillons de vins en fermentation alcoolique .....	71
c)	Echantillons de vins durant le vieillissement en barriques .....	71
d)	Echantillons de vins finis commerciaux .....	72
e)	Prétraitement des échantillons avant analyses .....	75
B.	Analyses spectroscopiques.....	75
1.	Mesure des spectres UV-visible.....	75
2.	Mesure des spectres IRTF.....	76
C.	Analyses de références des anthocyanes et des tanins.....	78
1.	Maturité phénolique déterminée par la méthode Glories.....	79
2.	Dosage des anthocyanes libres par décoloration au bisulfite de sodium .....	81

3.	Quantification et caractérisation des anthocyanes moléculaires par chromatographie liquide haute performance (CLHP) et détection UV-visible.....	81
4.	Détermination du ratio pigments monomériques/pigments polymérisés et de la concentration en tanins précipitables par la méthode de dosage d'Adams-Harbertson.....	82
5.	Détermination de la concentration en tanins par précipitation à la méthylcellulose .....	84
6.	Détermination des tanins totaux par hydrolyse acide (dosage de Bate-Smith).....	85
7.	Détermination du degré de polymérisation moyen (DPM) par phloroglucinolyse .....	85
a)	Principe général de la réaction de phloroglucinolyse .....	85
b)	Purification de l'échantillon.....	87
c)	Réaction de phloroglucinolyse et analyse en CLHP-UV .....	88
D.	Analyses sensorielles .....	89
1.	Entraînement des dégustateurs.....	89
2.	Dégustation des vins de l'étude .....	90
E.	Analyses des données par régression PLS .....	92
1.	Prétraitement des spectres.....	92
2.	Conception de modèles de prédiction par régression PLS.....	94
a)	Logiciels et algorithmes .....	94
b)	Choix de la validation croisée et du jeu de données pour la validation .....	94
c)	Sélection du nombre de variables latentes .....	95
<b>III.</b>	<b>Résultats et discussions .....</b>	<b>97</b>
A.	Prédiction des concentrations en anthocyanes et tanins du raisin pendant la maturité.....	99
1.	Analyse du jeu de données.....	100
a)	Evaluation de la variabilité des analyses de références .....	100
b)	Analyse en composantes principales (ACP) des analyses de références et des données spectrales.....	103
(1)	ACP des variables des analyses de références .....	103

(2) ACP des individus en fonction des analyses de références et des spectres UV-visible et IRTF .....	106
c) Analyse discriminante PLS pour la différenciation des millésimes étudiés pendant la maturité du raisin .....	108
2. Construction de modèles PLS pour la prédiction des anthocyanes et des tanins pendant la maturité du raisin .....	111
a) Construction de modèles de prédiction PLS pour les analyses Glories sur les millésimes 2017, 2018 et 2019 pendant la maturité du raisin .....	112
(1) Evolution des paramètres Glories pendant la maturité du raisin.....	112
(2) Conception de modèles de prédiction PLS pour les paramètres de la méthode Glories pendant la maturité du raisin.....	114
(3) Essai de validation des modèles de prédiction PLS en fonction du millésime pendant la maturité du raisin.....	117
b) Dosage des anthocyanes moléculaires sur les millésimes 2018 et 2019 pendant la maturité du raisin .....	120
(1) Evolution des anthocyanes moléculaires pendant la maturité du raisin.....	120
(2) Conception et validation de modèles de prédiction PLS pour le dosage des anthocyanes moléculaires pendant la maturité du raisin.....	123
Modèles de prédiction PLS du dosage des anthocyanes moléculaires par IRTF pendant la maturité du raisin .....	123
Modèles de prédiction PLS du dosage des anthocyanes moléculaires par UV-visible pendant la maturité du raisin .....	129
Modèles de prédiction PLS du dosage des anthocyanes moléculaires par IRTF-UV-visible pendant la maturité du raisin .....	131
c) Construction de modèles de prédiction PLS pour le suivi du ratio pigments monomériques/pigments polymériques sur les millésimes 2018 et 2019 pendant la maturité du raisin.....	133
d) Construction de modèles de prédiction PLS pour le suivi de l'évolution des tanins sur les millésimes 2018 et 2019 pendant la maturité du raisin .....	136
(1) Evolution des tanins pendant la maturité du raisin.....	136

(2) Conception et validation de modèles de prédiction PLS pour le dosage des tanins	138
3. Conclusions et perspectives de la conception de modèles de prédiction PLS pour l'analyse des anthocyanes et tanins du raisin pendant la maturité du raisin .....	140
B. Prédiction des concentrations en anthocyanes et tanins du moût et du vin rouge pendant la vinification .....	141
1. Analyse du jeu de données.....	141
a) Evaluation de la variabilité des analyses de références .....	141
b) Analyse en composantes principales (ACP) des analyses de références et des données spectrales.....	145
(1) ACP des variables des analyses de références .....	145
(2) ACP des individus en fonction des analyses de références et des spectres UV-visible et IRTF.....	147
c) Analyse discriminante PLS pour la différenciation des millésimes et des cépages étudiés .....	150
2. Evolution des concentrations en anthocyanes et en tanins pendant la fermentation alcoolique.....	152
a) Evolution des concentrations en anthocyanes pendant la fermentation alcoolique ...	153
b) Evolution de la composition en anthocyanes moléculaires pendant la fermentation alcoolique.....	155
c) Evolution de la polymérisation des anthocyanes pendant la fermentation alcoolique	157
d) Evolution des concentrations en tanins pendant la fermentation alcoolique .....	160
e) Evolution de la composition en unités flavan-3-ols des tanins pendant la fermentation alcoolique.....	162
f) Evolution du degré de polymérisation moyen des tanins pendant la fermentation alcoolique.....	164

3.	Construction de modèles de prédiction PLS pour le suivi des anthocyanes et des tanins pendant la vinification.....	166
a)	Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification .....	167
(1)	Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification par IRTF .....	167
(2)	Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification par UV-visible .....	172
(3)	Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification par combinaison de l'IRTF et l'UV-visible 175	
b)	Construction de modèles de prédiction PLS avec validation externe pour le suivi des anthocyanes et des tanins pendant la vinification des moûts de raisin .....	177
(1)	Construction de modèles de prédiction PLS avec validation par millésime pour le suivi des anthocyanes et des tanins pendant la vinification des moûts de raisins par IRTF 177	
(2)	Construction de modèles de prédiction PLS par IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone .....	181
(3)	Construction de modèles de prédiction PLS avec validation par millésime pour le suivi des anthocyanes et des tanins pendant la vinification par UV-visible .....	187
(4)	Construction de modèles de prédiction PLS par IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone .....	190
4.	Conclusions et perspectives de la conception de modèles de prédiction PLS pour l'analyse des anthocyanes et tanins du raisin pendant la fermentation .....	194
C.	Suivi de l'évolution des anthocyanes et des tanins pendant le vieillissement en barriques par IRTF et UV-visible couplé à l'analyse PLS.....	199
1.	Comparaison de l'évolution des anthocyanes et des tanins pendant l'élevage en barriques et en bouteilles .....	199
a)	Évolution des anthocyanes pendant l'élevage en barriques et en bouteilles .....	199

b)	Évolution des tanins pendant l'élevage en barriques et en bouteilles .....	202
2.	Discrimination du vieillissement en barriques et en bouteilles par IRTF et UV-visible .	204
3.	Prédiction des concentrations en anthocyanes et tanins de vins de Merlot pendant l'élevage à l'aide des modèles construits pour la fermentation .....	208
a)	Conclusions et perspectives de la prédiction des anthocyanes et des tanins par IRTF et UV-visible pendant l'élevage.....	210
D.	Construction de modèles de prédiction PLS pour le suivi des anthocyanes et des tanins dans les vins finis commerciaux.....	210
1.	Analyse du jeu de données.....	211
a)	Evaluation de la variabilité des analyses de références .....	211
b)	Analyse en composantes principales (ACP) des analyses de références et des données spectrales .....	216
2.	Construction de modèles de prédiction PLS pour le suivi des anthocyanes et des tanins sur les vins finis commerciaux.....	219
a)	Comparaison entre les modèles PLS construits avec les spectres IRTF et UV-visible pour la prédiction des anthocyanes et des tanins sur les vins finis commerciaux .....	219
b)	Combinaison des spectres IRTF et UV-visible pour la prédiction des anthocyanes et des tanins sur les vins finis commerciaux .....	223
3.	Modélisation de l'astringence et l'amertume par IRTF et UV-visible couplé à l'analyse PLS	227
a)	Prédiction de l'astringence et de l'amertume par IRTF couplé à l'analyse PLS .....	229
b)	Prédiction de l'astringence et de l'amertume par UV-visible couplé à l'analyse PLS	231
4.	Conclusions et perspectives de la prédiction des anthocyanes et des tanins par IRTF et UV-visible pendant l'élevage.....	232
E.	Modélisation globale des anthocyanes et tanins des moûts et vins par IRTF et UV-visible couplé à une analyse de régression PLS .....	236

1. Modélisation globale des anthocyanes et tanins des moûts et vins par IRTF couplé à une analyse de régression PLS.....	237
2. Modélisation globale des anthocyanes et tanins des moûts et vins par UV-visible couplé à une analyse de régression PLS.....	239
<b>Conclusion générale et perspectives .....</b>	<b>241</b>
<b>Références bibliographiques .....</b>	<b>253</b>

## Liste des figures

Figure 1 : Structure des stilbènes du raisin.....	13
Figure 2 : Structure générale des composés flavonoïdes.....	15
Figure 3 : (1) structure générale des flavones, (2) lutéoline, (3) génistéine .....	15
Figure 4 : Structures des flavanones retrouvés dans les vins : (1) naringinine, (2) hespéretine.....	16
Figure 5 : Structures des principaux flavonols aglycones : (1) kaempférol, (2) quercétine, (3) myricétine, (4) isorhamnétine .....	17
Figure 6 : Structure de la taxifoline .....	18
Figure 7 : Structures des principaux Flavan-3-ols du vin : (1) (-)-épicatéchine, (2) (-)-épigallocatechine, (3) (-)-épicatéchine-3- <i>O</i> -gallate, (4) (+)-catéchine, (5) (+)-gallocatechine .....	19
Figure 8 : Différents types de liaisons composants les tanins condensés .....	21
Figure 9 : Structure d'un tétramère de procyanidine cyclique .....	22
Figure 10 : Evolution des différentes formes des anthocyanes en fonction du pH .....	25
Figure 11 : Equilibre entre les différentes formes anthocyaniques : l'ion flavylum, la base quinonique, la base carbinol et la chalcone, exemple de la malvidine-3- <i>O</i> -glucoside.....	26
Figure 12 : Structures des produits de la réaction entre l'ion hydrogénosulfite et les anthocyanes .....	27
Figure 13 : (1) 1,2,3,4,6-penta- <i>O</i> -galloyl- $\beta$ -D-glucopyranose, précurseur des tanins hydrolysables ; (2) exemple de gallotanins, le 2- <i>O</i> -digalloyl-1,2,3,4,6-tétra- <i>O</i> -galloyl- $\beta$ -D-glucopyranose .....	28
Figure 14 : Structures des 8 ellagitannins C-glycosidiques principaux présents dans le bois de chêne.....	29
Figure 15 : Evolution de la teneur en ellagitannins C-glycosidiques durant l'élevage en barriques de chêne (163) .....	30
Figure 16 : Structure des principaux flavano-ellagitannins du vin.....	31
Figure 17 : Exemple de formation d'un anthocyano-ellagitannin, réaction de la vescalagine avec la malvidine-3- <i>O</i> -glucoside .....	32
Figure 18 : Polymérisation directe par catalyse acide : réaction flavan-3-ol/flavan-3-ol et flavan-3-ol/anthocyane.....	33
Figure 19 : Polymérisation directe et formation de liaison de type A, exemple de monomères et dimères d'épicatéchine.....	34
Figure 20 : Exemple de condensation directe entre la malvidine-3- <i>O</i> -glucoside et l'épicatéchine.....	35
Figure 21 : Polymérisation indirecte en présence d'acétaldéhyde : réaction flavan-3-ol/flavan-3-ol et flavan-3-ol/anthocyane .....	36
Figure 22 : Réaction de formations des pyranoanthocyanes de type Vitisines, exemple de la malvidine-3- <i>O</i> -glucoside, d'après de Freitas & Mateus, 2011.....	39

Figure 23 : Différentes structures des pyranoanthocyanes des vins : (1) exemple d'hydroxypyrananthocyanes, la pinotine A ; (2) flavano-pyranoanthocyane issue de la réaction entre la malvidine-3- <i>O</i> -glucoside et l'épicatéchine ; (3) dimère de pyranoanthocyane issu de la réaction entre deux malvidines-3- <i>O</i> -glucoside .....	39
Figure 24 : Schéma des interactions entre tanins condensés et protéines.....	42
Figure 25 : Intensité de la perception d'astringence et d'amertume en fonction du degré de polymérisation moyen des tanins condensés (265) .....	44
Figure 26 : Spectres de différents composés phénoliques .....	47
Figure 27 : Différents types de vibrations induits par un rayonnement infrarouge .....	50
Figure 28 : Spectre infrarouge moyen des tanins du vin .....	51
Figure 29 : Modèle hiérarchique pour une analyse PLS de J blocs de variables.....	55
Figure 30 : Représentation schématique de la validation croisée <i>leave-one-out</i> .....	58
Figure 31 : Répartition des millésimes du panel d'échantillons de vins finis commerciaux.....	74
Figure 32 : Répartition des cépages principaux du panel d'échantillons de vins finis commerciaux .....	74
Figure 33 : Spectres d'absorbance UV-visible du panel de vins finis commerciaux .....	76
Figure 34 : (1) Spectres IRTF complets du panel de vins finis commerciaux, (2) Spectres IRTF avec les longueurs d'ondes sélectionnées du panel de vins finis commerciaux.....	78
Figure 35 : Etape de filtration de la méthode Glories, obtention des filtrats qui permettront l'analyse de la composition phénolique.....	80
Figure 36 : Dosages des pigments et tanins précipitables (275).....	84
Figure 37 : Réaction de phloroglucinolyse sur les tanins condensés (277) .....	87
Figure 38 : Poste de dégustation type proposé pour définir l'astringence et l'amertume des vins étudiés ....	91
Figure 39 : Feuille de dégustation type proposée pour définir l'astringence et l'amertume des vins étudiés .	92
Figure 40 : (1) Spectres UV-visible et (2) spectres IRTF du panel de vins finis commerciaux après centrage par la moyenne .....	93
Figure 41 : Erreur quadratique moyenne de calibration (RMSEC) et de validation croisée (RMSECV) en fonction du nombre de variables latentes : exemple de modèles IRTF sur un panel de vins finis commerciaux. ....	95
Figure 42 : ACP des variables des analyses de références : (1) dimensions 1 et 2, (2) dimensions 3 et 4 (RPT : richesse polyphénolique totale.....	104
Figure 43 : ACP des variables des anthocyanes moléculaires, dimension 1 et 2 .....	105
Figure 44 : ACP des individus des analyses de références : (1) dimensions 1 et 2, (2) dimensions 3 et 4 ...	106
Figure 45 : ACP des individus des spectres UV-visible.....	107
Figure 46 : ACP des individus des spectres IRTF.....	108
Figure 47 : Analyse PLS-DA des millésimes en fonction des analyses de références .....	109

Figure 48 : Analyse PLS-DA des millésimes en fonction des spectres UV-visible et IRTF.....	110
Figure 49 : Evolution du potentiel total en anthocyanes (ApH1), du potentiel en anthocyanes extractibles (ApH3,2), et de la richesse polyphénolique totale (RPT) sur les millésimes 2017, 2018 et 2019 .....	113
Figure 50 : ApH3,2 mesuré en fonction d'ApH3,2 prédit par les modèles de prédiction PLS de l'IRTF. Séparation de la calibration et de la validation en fonction des millésimes. ....	119
Figure 51 : Evolution de la concentration en anthocyanes moléculaires sur les millésimes 2018 et 2019 (M : Merlot, Cf : Cabernet franc) .....	121
Figure 52 : Distribution des anthocyanes en fonction des parcelles étudiées.....	122
Figure 53 : Prédiction de la concentration en pétonidine-3- <i>O</i> -acétylglucoside en fonction du modèle de prédiction des anthocyanes totales par IRTF.....	126
Figure 54 : Corrélacion entre les concentrations en pétonidine-3- <i>O</i> -acétylglucoside et en anthocyanes totales .....	126
Figure 55 : Prédiction de la concentration modifiée de pétonidine-3- <i>O</i> -acétylglucoside en fonction du modèle de prédiction des anthocyanes totales par IRTF.....	127
Figure 56 : Suivi du ratio pigments monomériques (MP), petits pigments polymérisés (SPP) et grands pigments polymérisés (LPP) pendant la maturité.....	134
Figure 57 : Evolution de la concentration en tanins sur les millésimes 2018 et 2019 par dosage à l'albumine de sérum bovin et à la méthylcellulose.....	137
Figure 58 : ACP des variables des analyses de références des moûts et vins en fermentation. (1) ACP de tous les paramètres à l'exception des concentrations en anthocyanes moléculaires individuelles, (2) ACP des concentrations anthocyanes moléculaires individuelles et des sommes des anthocyanes.....	146
Figure 59 : ACP des individus des analyses de références : (1) représentation en fonction du cépage (2) représentation en fonction du millésime.....	147
Figure 60 : ACP des individus des spectres UV-visible : (1) représentation en fonction du cépage (2) représentation en fonction du millésime.....	148
Figure 61 : ACP des individus des spectres IRTF : (1) représentation en fonction du cépage (2) représentation en fonction du millésime.....	149
Figure 62 : ACP des individus des spectres IRTF : représentation en fonction du temps de fermentation ..	150
Figure 63 : Evolution de la concentration en anthocyanes libres et anthocyanes moléculaires sur les millésimes 2017, 2018 et 2019 .....	153
Figure 64 : Distribution des différentes anthocyanes moléculaires en fonction des cuves de vinification étudiées pour les millésimes 2017, 2018 et 2019 .....	155
Figure 65 : Evolution de la concentration et la distribution des anthocyanes moléculaires de la cuve M11 durant la fermentation alcoolique .....	156

Figure 66 : Evolution de la polymérisation des anthocyanes en fonction du temps de fermentation pour les millésimes 2017, 2018 et 2019 sur les vins en fermentation.....	158
Figure 67 : Evolution des concentrations en tanins dosés par précipitation à l'albumine de sérum bovin (BSA) et par précipitation à la méthylcellulose (MCP) en fonction du temps de fermentation pour les cuves des millésimes 2017, 2018 et 2019.....	160
Figure 68 : Evolution de la composition en unités flavan-3-ols en fonction du temps de fermentation pour les millésimes 2017, 2018 et 2019 sur les vins en fermentation.....	163
Figure 69 : Evolution du degré de polymérisation moyen en fonction du temps de fermentation pour les millésimes 2017, 2018 et 2019 pour les vins en fermentation .....	165
Figure 70 : modèle de prédiction de la somme des anthocyanes moléculaires avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation .....	184
Figure 71 : modèles de prédiction des tanins précipitables à l'albumine de sérum bovin et la méthylcellulose avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation.....	186
Figure 72 : modèle de prédiction des anthocyanes libres avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation.....	192
Figure 73 : modèles de prédiction des tanins précipitables à l'albumine de sérum bovin et la méthylcellulose avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation.....	193
Figure 74 : récapitulatif de la modélisation des anthocyanes et des tanins pendant la fermentation par IRTF et UV-visible .....	198
Figure 75 : Evolution de la concentration et de la composition en anthocyanes moléculaires de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles.....	200
Figure 76 : Evolution de la proportion en pigments monomériques et polymériques de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles.....	201
Figure 77 : Evolution de la concentration en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles .....	202
Figure 78 : Evolution du degré de polymérisation moyen de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles .....	203
Figure 79 : Analyse PLSDA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des spectres UV-visible .....	204
Figure 80 : Analyse PLSDA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des spectres IRTF.....	205
Figure 81 : Analyse PLSDA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des spectres IRTF-UV-visible .....	206

Figure 82 : Analyse PLS-DA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des analyses en anthocyanes moléculaires, pigments monomériques et polymériques.....	207
Figure 83 : Prédiction des anthocyanes libres durant l'élevage pour les vins de Merlot étudiés par le modèle UV-visible construit avec les échantillons de moûts et vins du suivi de la fermentation.....	209
Figure 84 : Concentration de la somme des anthocyanes moléculaires en fonction du millésime sur les vins finis commerciaux .....	213
Figure 85 : Proportion en pigments monomériques, petits pigments polymérisés et grands pigments polymérisés en fonction du millésime sur les vins finis commerciaux .....	214
Figure 86 : Evolution du degré de polymérisation moyen en fonction du millésime sur les vins finis commerciaux .....	215
Figure 87 : Analyses en composantes principales, dimension 1 et 2 : (1) individus des analyses de références, (2) variables des analyses de références, (3) individus des spectres IRTF, (4) individus des spectres UV-visible .....	216
Figure 88 : Analyses en composantes principales des individus des analyses de références en fonction des millésimes (1) et des cépages (2), dimension 1 et 2 .....	218
Figure 89 : Analyse en composantes principales des variables des vins dégustés, avec astringence et amertume en variables externes.....	228
Figure 90 : Modèle de prédiction de l'astringence et l'amertume mesurées par le panel de dégustateurs par IRTF couplé à l'analyse PLS.....	230
Figure 91 : Modèle de prédiction de l'astringence et l'amertume mesurées par le panel de dégustateurs par UV-visible couplé à l'analyse PLS.....	231
Figure 92 : récapitulatif de la modélisation des anthocyanes et des tanins sur les vins finis commerciaux par IRTF et UV-visible.....	235
Figure 93 : Méthodologie de conception et validation de modèles de régression PLS grâce à l'IRTF et l'UV-visible pour la prédiction des concentrations et compositions en anthocyanes et tanins des raisins, moûts et vins. ....	244
Figure 94 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour le suivi de la maturité du raisin .....	246
Figure 95 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour le suivi de la fermentation alcoolique des vins rouges.....	248
Figure 96 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour le suivi des vins finis commerciaux .....	250
Figure 97 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour la modélisation globale (fermentation, élevage et vins commerciaux).....	252

## Liste des tableaux

Tableau 1 : Structure des acides hydroxybenzoïques du vin.....	11
Tableau 2 : Structure des acides hydroxycinnamiques du vin.....	12
Tableau 3 : Structures des principales anthocyanes du raisin et du vin.....	24
Tableau 4 : Absorbance à 280 nm de plusieurs polyphénols du vin, dilué à 1 mg/L dans une solution à 10 % de méthanol et 1 % d'acide chlorhydrique.....	48
Tableau 5 : Longueurs d'ondes maximales d'absorbance et coefficients d'extinction molaire des cinq principales anthocyanes monoglucosidées du vin, en solution aqueuse à 10 % d'éthanol et pH 1,5.....	49
Tableau 6 : Principales réponses spectrales des groupes fonctionnels composant les polyphénols du vin.....	52
Tableau 7 : Solvants, acides, bases, réactifs et standards utilisés.....	69
Tableau 8 : Millésime, origine et cépages des vins finis commerciaux de l'étude (les vins surlignés ont été testés par le panel de dégustateurs entraînés.....	73
Tableau 9 : Gradient d'élution pour la séparation des anthocyanes moléculaires en CLHP.....	82
Tableau 10 : Gradient d'élution pour la séparation des flavan-3-ols et flavan-3-ol-phloroglucinol en CLHP.....	88
Tableau 11 : Masse moléculaire d'ionisation des flavan-3-ols et flavan-3-ol-phloroglucinol.....	88
Tableau 12 : Paramètres des analyses de références obtenus sur les échantillons de raisins pendant la maturité.....	101
Tableau 13 : Modèles de prédiction PLS des paramètres de la méthode Glories par UV-visible, IRTF et IRTF-UV-visible. Utilisation du jeu de données complet pour la calibration.....	115
Tableau 14 : Résultats de prédiction des modèles PLS des paramètres de la méthode Glories en fonction du millésime.....	118
Tableau 15 : Modèles de prédiction PLS du dosage des anthocyanes moléculaires par IRTF.....	124
Tableau 16 : Modèles de prédiction PLS du dosage des anthocyanes moléculaires par UV-visible.....	130
Tableau 17 : Modèles de prédiction PLS du dosage des anthocyanes totales par IRTF-UV-visible.....	132
Tableau 18 : Modèles de prédiction PLS du dosage des pigments monomériques (MP), petits pigments polymérisés (SPP) grands pigments polymérisés (LPP) et pigments polymérisés totaux (PP = SPP + LPP).....	135
Tableau 19 : Modèles de prédiction PLS pour les tanins précipitables à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP) par UV-visible, IRTF et IRTF-UV-visible. Utilisation du jeu de données complet pour la calibration.....	138
Tableau 20 : Résultats de prédiction PLS avec validation par échantillons externes pour les tanins précipitables à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP) par UV-visible, IRTF et IRTF-UV-visible.....	139
Tableau 21 : Paramètres des analyses de références obtenus sur les échantillons de moûts et vins en fermentation.....	143

Tableau 22 : Analyses PLS discriminantes pour la différenciation des millésimes et des cépages du jeu de données .....	151
Tableau 23 : Modèles de prédiction PLS par IRTF avec validation croisée pour tous les paramètres d'anthocyanes et de tanins suivis pendant la fermentation.....	168
Tableau 24 : Corrélation des différentes anthocyanes avec la somme des anthocyanes moléculaires et résultats de prédictions des différentes anthocyanes avec le modèle de prédiction de la somme des anthocyanes moléculaires.....	170
Tableau 25 : Modèles de prédiction PLS par UV-visible avec validation croisée pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation des raisins .....	173
Tableau 26 : Modèles de prédiction PLS par IRTF-UV-visible avec validation croisée pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation.....	176
Tableau 27 : Modèles de prédiction PLS par IRTF avec validation par millésimes pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation.....	179
Tableau 28 : modèles de prédiction PLS par IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation .....	182
Tableau 29 : Modèles de prédiction PLS par UV-visible avec validation par millésimes pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation.....	188
Tableau 30 : modèles de prédiction PLS par UV-visible avec validation externe déterminée par l'algorithme de Kennard-Stone pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation de moûts de raisins.....	191
Tableau 31 : Prédiction des anthocyanes et des tanins durant l'élevage pour les vins de Merlot étudiés par les modèles IRTF et UV-visible construits avec les échantillons de moûts et vins du suivi de la fermentation	208
Tableau 32 : Paramètres des analyses de références obtenus sur les vins finis commerciaux .....	212
Tableau 33 : Prédiction des anthocyanes et des tanins sur les vins commerciaux par les modèles IRTF, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone.....	220
Tableau 34 : Prédiction des anthocyanes et des tanins sur les vins commerciaux par les modèles UV-visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone.....	221
Tableau 35 : Prédiction des anthocyanes et des tanins sur les vins par les modèles IRTF-UV-visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone.....	224
Tableau 36 : Prédiction des anthocyanes sur les vins par les modèles IRTF combinés aux longueurs d'ondes 420, 520 et 620 nm du visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone.....	226
Tableau 37 : Coefficients de régression linéaire des paramètres décrivant les tanins en fonction de l'astringence mesurée par le panel de dégustateurs .....	228

Tableau 38 : Prédiction des anthocyanes et des tanins sur les vins en fermentation, en élevage et commerciaux par les modèles IRTF, validation croisée et validation externe déterminée par l’algorithme de Kennard-Stone .....	237
Tableau 39 : Prédiction des pigments monomériques et polymériques sur les vins en fermentation, en élevage et commerciaux par les modèles IRTF avec l’ajout des longueurs d’ondes du visible 420, 520 et 620 nm, validation croisée et validation externe déterminée par l’algorithme de Kennard-Stone .....	238
Tableau 40 : Prédiction des anthocyanes et des tanins sur les vins en fermentation, en élevage et commerciaux par les modèles UV-visible, validation croisée et validation externe déterminée par l’algorithme de Kennard-Stone .....	239

## Liste des abréviations

**ACP** : Analyse en composantes principales

**BSA** : *Bovin serum albumin*, Sérum d'albumine bovin

**CLHP** : Chromatographie liquide haute performance

**Cy-3-O-acglc** : Cyanidine-3-O-acétylglucoside

**Cy-3-O-glc** : Cyanidine-3-O-glucoside

**Dp-3-O-acglc** : Delphinidine-3-O-acétylglucoside

**Dp-3-O-glc** : Delphinidine-3-O-glucoside

**DPm** : Degré de polymérisation moyen

**LPP** : *Large polymeric pigments*, grands pigments polymérisés

**MCP** : *Methylcellulose polymer*, polymère de méthylcellulose

**Mlv-3-O-acglc** : Malvidine-3-O-acétylglucoside

**Mlv-3-O-cmglc** : Malvidine-3-O-coumaroylglucoside

**Mlv-3-O-glc** : Malvidine-3-O-glucoside

**MP** : *Monomeric pigments*, pigments monomériques

**PLS** : *Partial least squares regression*, régression des moindres carrés partiels

**PLSDA** : *Partial least squares discriminant analysis*, analyse des moindres carrés partiels discriminante

**Pn-3-O-acglc** : Péonidine-3-O-acétylglucoside

**Pn-3-O-cmglc** : Péonidine-3-O-coumaroylglucoside

**Pn-3-O-glc** : Péonidine-3-O-glucoside

**Pt-3-O-acglc** : Pétunidine-3-O-acétylglucoside

**Pt-3-O-glc** : Pétunidine-3-O-glucoside

**R<sup>2</sup>** : Coefficient de régression linéaire

**RMSE** : *Root mean square error*, erreur quadratique moyenne

**RPD** : *residual predictive deviation*, écart prédictif résiduel

**SPP** : *small polymeric pigments*, petits pigments polymérisés

**VL** : variable latente



---

# **Introduction et objectifs**

---



Dans les vins rouges, les composés polyphénoliques extraits du raisin sont présents en concentration importante. La famille des flavonoïdes, en particulier, a un impact significatif sur la qualité des vins ainsi que sur leur vieillissement (1). Les perceptions sensorielles comme l'astringence et l'amertume sont principalement liées à la concentration et la composition en tanins condensés, tandis que l'intensité et l'évolution de la couleur sont dues à la concentration en anthocyanes (2–5). Au cours du vieillissement ces molécules évoluent : elles se dégradent, polymérisent, et impactent ainsi les propriétés organoleptiques du vin. Afin d'optimiser et obtenir le produit final avec le profil analytique et sensoriel recherché, il apparaît alors essentiel pour le vinificateur de pouvoir mesurer la concentration et la composition en anthocyanes et en tanins dans le raisin, le moût et le vin.

Plusieurs analyses ont été développées pour mesurer les concentrations en polyphénols. Cependant, la plupart de ces méthodes nécessitent du matériel de laboratoire, des connaissances et du temps, et ne sont donc pas adaptées à une analyse rapide de routine dans le milieu industriel. Pour faciliter l'analyse de ces composés, il est alors possible de développer des méthodes de spectroscopie couplées à la chimométrie. L'empreinte spectrale d'un moût ou d'un vin apporte un nombre important d'informations sur la composition en polyphénols, mais ces données peuvent être compliquées à extraire. Le moût et le vin sont des matrices complexes, et de nombreux composés absorbent aux longueurs d'onde d'intérêts des anthocyanes et des tanins, ne permettant pas une lecture directe de leur concentration. De ce fait, l'analyse chimométrique couplée à l'analyse spectrale apparaît nécessaire pour concevoir des modèles de prédiction pour les composés recherchés (6). Si différentes analyses statistiques ont montré des résultats concluants, la régression partielle des moindres carrés (PLS) apparaît être la plus efficace pour produire des modèles de prédiction robustes et efficaces (7,8).

Plusieurs méthodes spectroscopiques couplées à une analyse de régression PLS ont déjà démontrées leur potentiel pour l'analyse des anthocyanes et des tanins présents dans les moûts et les vins. Deux analyses spectrales se sont avérées particulièrement efficaces : l'ultraviolet-visible (UV-visible) et l'infrarouge à transformée de Fourier (IRTF). Plusieurs études ont ainsi démontré l'efficacité de ces deux méthodes spectroscopiques couplées à une analyse multivariée PLS pour développer des modèles de prédiction de la concentration en anthocyanes et en tanins (9–15). Cependant, certains de ces travaux sont limités en termes de nombre d'échantillons, de diversité, ou de robustesse de la prédiction. De plus, les limites de prédiction ainsi que les biais possibles dus à une surinterprétation des résultats ne sont jamais développés. Par ailleurs, il n'existe aucun travaux de recherche à notre connaissance comparant les deux méthodes spectroscopiques UV-visible et IRTF, et la

complémentarité possible des deux spectres n'a jamais été recherchée, et aucune étude à ce jour n'a démontré la possibilité de modéliser directement des perceptions comme l'astringence et l'amertume grâce à une analyse spectrale couplée à l'analyse de régression PLS.

Notre travail de thèse a donc pour objectif principal d'identifier le potentiel de la spectroscopie UV-visible et IRTF couplée à la régression PLS pour la prédiction des concentrations en anthocyanes et tanins du raisin, du moût et du vin rouge. Notre approche cherche à approfondir les avantages et les limites de ces techniques analytiques dans le but de développer des modèles performants pour la quantification des anthocyanes et des tanins applicables en milieu industriel viti-vinicole (production et contrôle). De plus, nous avons cherché à étudier la possibilité de modéliser des paramètres organoleptiques tels que l'astringence ou l'amertume grâce aux analyses spectrales.

Nos objectifs et plan de recherche sont présentés dans plusieurs volets :

1. Sélection d'échantillons de raisins, moûts et vins : ceux-ci doivent être représentatifs des étapes de vinification et de vieillissement. Le matériel utilisé pour la calibration des modèles est primordial et doit refléter la population d'échantillons qui sera analysée en routine.
2. Analyse spectrale : elle doit être la plus simple possible et limiter au maximum la manipulation des échantillons. Le principal intérêt recherché est que celle-ci puisse être facilement applicable, pour permettre une analyse rapide.
3. Analyses de références : elles doivent être représentatives de la composition de la matrice étudiée. Il est nécessaire de choisir des analyses qui permettront de suivre l'évolution des anthocyanes et des tanins condensés dans leur globalité, mais aussi d'étudier leur composition pour mieux comprendre leur évolution.
4. Analyses sensorielles : l'astringence et l'amertume sont deux paramètres importants pour le dégustateur et consommateur. Pour essayer d'appréhender ces propriétés, les vins devront être testés par un panel de dégustateurs, pour tenter de corréler la perception avec la composition phénolique ou directement avec la réponse spectrale.
5. Création de modèles de prédiction : en appliquant la régression PLS aux données spectrales et aux analyses de références, le but est de créer des modèles robustes capables d'expliquer les variations observées. Leur développement est une étape importante car indépendamment de la base de données utilisée, le choix des différents paramètres de la régression PLS tels que

le prétraitement des données ou le choix du nombre de variables va influencer grandement la prédictivité finale.

6. Validation des modèles obtenus : il est primordial pour une utilisation généralisée que les prédictions obtenues soient robustes et s'appliquent avec une erreur acceptable pour les échantillons extérieurs à la calibration.

Pour atteindre nos objectifs, nous présenterons dans un premier temps l'état des connaissances actuelles sur les polyphénols, leur réponse spectrale en UV-visible et IRTF, ainsi que l'analyse multivariée PLS appliquée aux anthocyanes et tanins du raisin, du moût et du vin. Dans un deuxième temps, le matériel, les échantillons et les méthodes utilisés pour concevoir les modèles de prédiction recherchés seront abordés et explicités. Enfin, la troisième partie de ce travail présentera les résultats obtenus, qui seront détaillés afin de discuter de la robustesse des prédictivités observées pour chaque paramètre étudié, et ainsi mettre en évidence les intérêts et les limites de la spectroscopie UV-visible et IRTF couplée à la régression PLS pour la prédiction des anthocyanes et tanins du raisin, du moût et du vin rouge.



---

# **I. Etude bibliographique**

---



## **A. Les polyphénols du vin : composition, évolution et impact sensoriel**

Les plantes sont une source inépuisable de diversité pour les composés organiques, avec de nombreuses familles de molécules. Parmi celles-ci il existe les polyphénols, métabolites secondaires les plus abondants du règne végétal (16). Ces composés se déclinent sous la forme de plusieurs milliers de structures différentes connues à ce jour, et apportent à la plante une protection contre les rayons ultraviolets, les pathogènes ou les parasites (17,18).

Omniprésents dans l'alimentation humaine, on les retrouve notamment dans un grand nombre d'aliments d'origine végétale (fruits, légumes, céréales, chocolat, etc) mais aussi dans un grand nombre de boissons (thé, café, bière, vin, etc) (19). Les polyphénols peuvent avoir un impact organoleptique et sensoriel important. Ainsi, dans les vins rouges, ils sont majoritairement responsables de l'astringence, l'amertume et la couleur (2–5).

Outre l'intérêt gustatif, les polyphénols sont recherchés dans l'alimentation pour leurs bienfaits sur la santé. Un grand nombre d'études épidémiologiques ont démontré les implications de ces composés dans la protection des maladies cardio-vasculaires (20–23), de plusieurs cancers (24–27), et des maladies neurodégénératives (28–30). Si le mécanisme primaire d'action des polyphénols a été attribué à leur effet antioxydant, celui-ci n'est plus considéré comme étant la source principale de bio protection car les concentrations relevées dans les tissus montrent des niveaux parfois trop faibles pour le piégeage des radicaux libres (31,32). Cependant, un certain nombre d'autres mécanismes biochimiques et moléculaires possibles ont été identifiés, notamment dans les voies de signalisation intracellulaires et intercellulaires, expliquant ainsi l'impact bénéfique de leur consommation sur la santé à des concentrations faibles (33,34).

Avant d'être appelé polyphénols, ces composés issus des plantes étaient nommés « tanins végétaux » en raison de leur utilisation dans le tannage des peaux animales. Du point de vue de la caractérisation moléculaire, si les phénols sont facilement identifiables par la présence d'un cycle benzénique associé à une fonction hydroxyle, la famille des polyphénols a vu sa définition évoluer au fil du temps. Le premier essai de caractérisation remonte à 1962 par Swain et Bate-Smith. Ils décrivent ainsi les polyphénols comme des phénols solubles dans l'eau, avec un poids moléculaire compris entre 500 et 3000 Da et avec la propriété de précipiter les alcaloïdes, la gélatine et diverses protéines (35). Cette définition a été approfondie en 1994 par Haslam et Cai, qui précisent que pour 1000 Da de masse

moléculaire relative, un polyphénol possède entre 12 et 16 groupements phénoliques et entre 5 et 7 cycles aromatiques (36). En se basant sur les caractéristiques évoquées précédemment, un concept plus précis a été proposé en 2011 par Quideau. Ainsi, le terme « polyphénol » ne devrait être utilisé que pour définir les métabolites secondaires de la plante dérivées exclusivement de la voie du shikimate-phénylpropanoïde et/ou de la voie des polycétides. De plus, ces composés doivent posséder plus d'un cycle phénolique et être dépourvus de tout groupe fonctionnel à base d'azote dans l'expression de leur structure la plus basique (37).

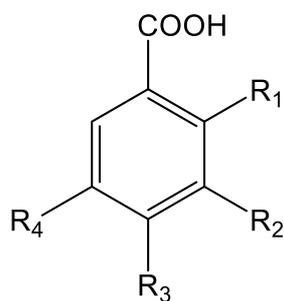
Lors de la vinification du vin rouge, les polyphénols sont extraits pendant la fermentation alcoolique, lorsque les parties solides des pépins et des pellicules sont en contact avec le jus de raisin et permettent leur solubilisation (38,39). Ensuite, leur structure est amenée à évoluer pendant l'élevage au contact du bois de la barrique et au cours du vieillissement (40,41). Au regard de la composition polyphénolique des vins rouges, il est possible de diviser cette famille en deux groupes, les composés flavonoïdes et les composés non-flavonoïdes (42), auxquels vient s'ajouter un troisième groupe, les tanins hydrolysables, issus de l'élevage des vins en présence de bois.

## **1. Les composés non-flavonoïdes**

Les composants non-flavonoïdes du vin sont constitués en majorité par les acides phénols et les stilbènes. Bien que non colorés, ils sont connus pour améliorer et stabiliser la couleur des vins rouges par des réactions intermoléculaires (43). Certains d'entre eux, comme le resvératrol, présentent une forte activité biologique et sont recherchés pour leurs bienfaits sur la santé (44,45).

### ***a) Les acides phénols***

Les acides phénols retrouvés dans le vin sont dérivés de deux structures types : l'acide hydroxybenzoïque et l'acide hydroxycinnamique. Les dérivés de l'acide hydroxybenzoïque s'articulent autour d'un squelette carboné de type C6-C1 (Tableau 1).

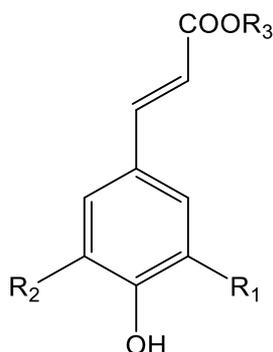


Acide hydroxybenzoïque	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>
Acide gallique	H	OH	OH	OH
Acide gentisique	OH	H	H	OH
Acide p-hydroxybenzoïque	H	H	OH	H
Acide protocatéchique	H	OH	OH	H
Acide salicylique	OH	H	H	H
Acide syringique	H	OCH <sub>3</sub>	OH	OCH <sub>3</sub>
Acide vanillique	H	OCH <sub>3</sub>	OH	H

Tableau 1 : Structure des acides hydroxybenzoïques du vin

Ces composants sont généralement trouvés sous leur forme libre dans le vin (46), avec comme principal représentant l'acide gallique dont les concentrations sont comprises entre 100 et 230 mg/kg (47). Son origine n'est pas seulement végétale par extraction du raisin, mais découle aussi de l'hydrolyse de certains tanins condensés et tanins hydrolysables, esters de l'acide gallique (48). Il a aussi été démontré que les acides hydroxybenzoïques pouvaient être le fruit de la dégradation des anthocyanes, sous l'action de la chaleur (49).

Les acides hydroxycinnamiques s'articulent autour d'un squelette carboné de type C6-C3 (Tableau 2).



Acide hydroxycinnamique	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>
Acide caféique	OH	H	H
Acide caftarique	OH	H	Acide tartrique*
Acide <i>p</i> -coumarique	H	H	H
Acide <i>p</i> -coutarique	H	H	Acide tartrique*
Acide férulique	OCH <sub>3</sub>	H	H
Acide fertarique	OCH <sub>3</sub>	H	Acide tartrique*
Acide sinapique	OCH <sub>3</sub>	OCH <sub>3</sub>	H

Tableau 2 : Structure des acides hydroxycinnamiques du vin (\*Structure de l'acide tartrique : HOOC-CHOH-CHOH-COOH)

Leur origine se situe dans la pulpe de la baie de raisin, principalement sous la forme d'esters de l'acide tartrique, les formes libres étant très minoritaires (50,51). Ces dérivés peuvent se trouver sous la configuration *cis*- et *trans*-, bien que la configuration *trans*-, plus stable, soit prédominante (52). Parmi les acides hydroxycinnamiques, l'acide caftarique participe jusqu'à 50 % du total de ces composés avec des concentrations de l'ordre de 200 mg/kg, viennent ensuite l'acide *p*-coumarique et l'acide fertarique (47,53).

Contrairement aux acides hydroxybenzoïques, ces composés ont un impact non négligeable sur le vin. La couleur du vin rouge peut ainsi être influencée par la présence des acides hydroxycinnamiques qui jouent un rôle important dans le phénomène de copigmentation, décrit par une intensification de la couleur des anthocyanes due à la présence d'autres composés phénoliques (43,54). Ces acides peuvent aussi réagir directement avec les anthocyanes pour former des pyranoanthocyanes (55,56). D'un point de vue organoleptique, la dégradation des acides *p*-coumariques et féruliques par des levures de type *Brettanomyces* ou *Dekkera* entraîne la formation de précurseurs des phénols volatils du vin, responsables d'un défaut olfactif caractérisé par une odeur de « sueur de cheval » ou « d'écurie » (57,58).

## b) Les stilbènes

Les stilbènes constituent une famille de polyphénols présente naturellement dans plusieurs types de plantes, mais le raisin et le vin sont considérés comme étant la source alimentaire la plus importante pour ces composés (59). Ils sont synthétisés par la plante pour répondre aux stress extérieurs comme les infections microbiennes ou les irradiations UV, et sont transférés du raisin au vin lors du processus de vinification (60,61). La structure principale des stilbènes est de type C6-C2-C6, et un nombre important de dérivés ont été identifiés dans le raisin (Figure 1) (62). Les stilbènes apparaissent sous la forme de monomères ou de polymères appelés viniférine. Cette polymérisation oxydative est induite par l'action d'une peroxydase sur le resvératrol (63).

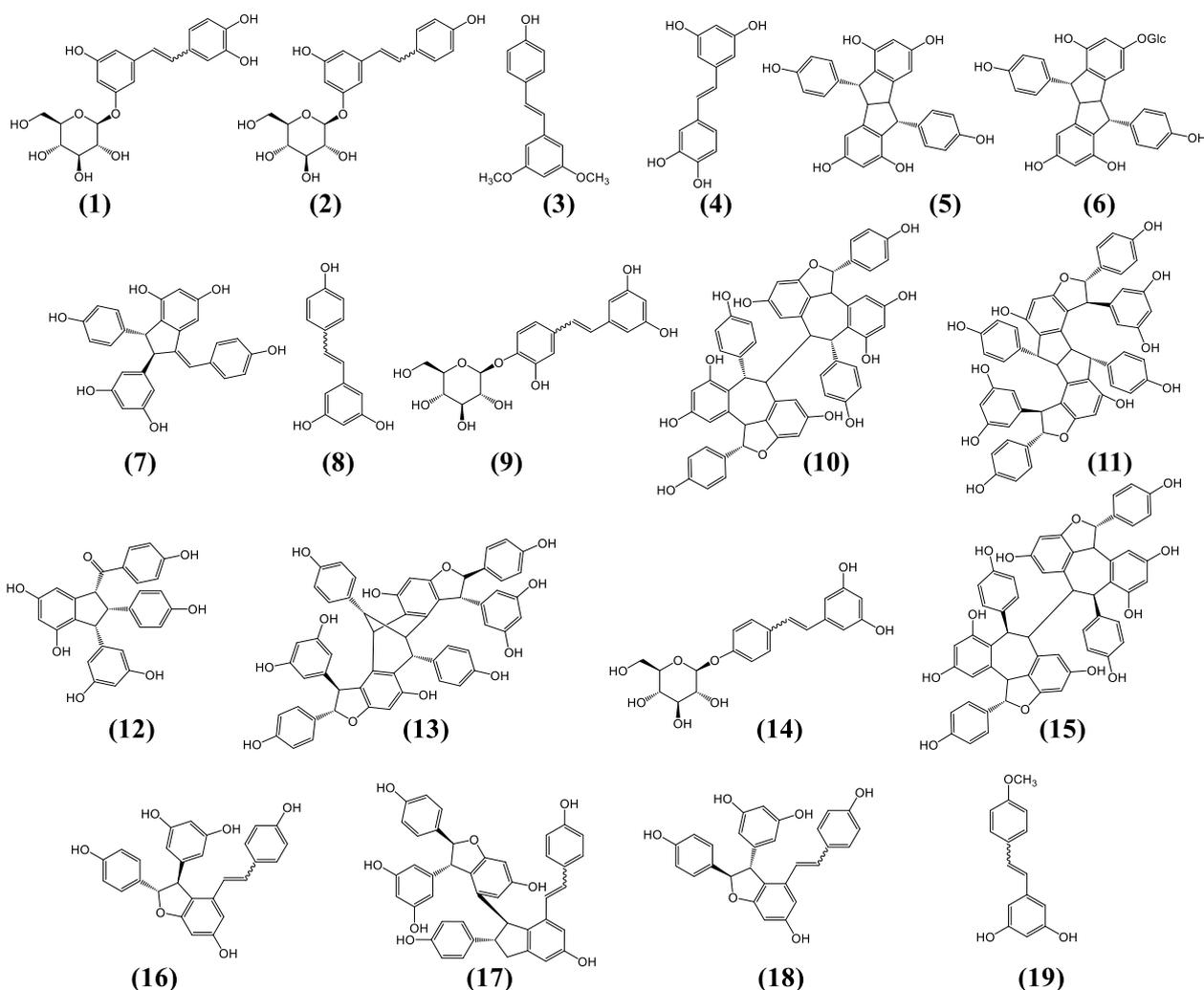


Figure 1 : Structure des stilbènes du raisin : (1) E- et Z-astringine ; (2) E- et Z-picéide ; (3) pterostilbène ; (4) picéatannol ; (5) pallidol ; (6) pallidol-3-O-glucoside ; (7) parthenocissine ; (8) *trans*- et *cis*-resvératrol ; (9) resvératroloside ; (10) hopéaphénol ; (11) ampelopsine H ; (12) caraphénol B ; (13) vaticanol C ; (14) resvératrol-4'-O-β-D-glucopyranoside ; (15) isohopéaphénol ; (16) E- et Z-ε-viniférine ; (17) E- et Z-miyabénol C ; (18) E- et Z-δ-viniférine ; (19) *trans*- et *cis*-resveratrol-4'-méthyl éther

La concentration en stilbènes d'un vin varie considérablement en fonctions de plusieurs facteurs incluant le climat, le cépage ou encore les infections fongiques (64–66). Il a notamment été démontré une corrélation positive entre la synthèse de *trans*-resvératrol dans la baie de raisin et le niveau d'infection relatif à *botrytis cinerea*, avec une production accrue localisée autour des lésions engendrées par l'attaque fongique (67). Cette synthèse a lieu particulièrement dans la pellicule du raisin, et explique la concentration en stilbènes supérieure dans les vins rouges en comparaison des vins rosés ou blancs, dont l'extraction des composés de la pellicule pendant la vinification est moins importante (65,68). Les stilbènes les plus représentés dans le raisin et le vin sont le resvératrol et son dérivé glycosylé, le picéide, qui participent pour plus de 50 % à la concentration en stilbène totale. En fonction des vins étudiés, la concentration en stilbènes dépasse rarement 30 mg/L, avec une moyenne de l'ordre de 10 mg/L (68,69). Ces composés sont très photosensibles, et bien que les configurations *trans*- ou *Z*- soient plus stables chimiquement, elles peuvent être isomérisés par irradiation UV et conduire aux formes *cis*- ou *E*- (70).

Les stilbènes, en particulier le *trans*-resvératrol, sont extrêmement recherchés pour leurs propriétés sur la santé humaine (71). Ils possèdent en effet des propriétés antioxydantes (72,73), cardioprotectrices (74,75), anticancéreuses (76,77) et anti-inflammatoires (78,79). Si ces molécules semblent particulièrement impliquées dans les effets bénéfiques du vin rouge sur la santé, il faut cependant être vigilant car de nombreuses études impliquent des doses journalières en resvératrol bien supérieures à la moyenne de celles amenées par une consommation de vin modérée (80,81).

## **2. Les composés flavonoïdes**

Les composés flavonoïdes constituent la famille de polyphénols la plus abondante dans les vins rouges (82). Ils ont un impact organoleptique important sur les vins rouges et contribuent majoritairement à l'astringence, l'amertume et la couleur (2–5). Ils se distinguent par une structure à 15 carbones, composée d'un squelette de type C6-C3-C6 consistant en deux cycles aromatiques A et B, liés ensemble par un hétérocycle C (Figure 2). Plusieurs sous-classes de ces composés se retrouvent dans les vins rouges : les flavones, les flavanones, les flavonols, les flavononols, les flavan-3-ols et les anthocyanes.

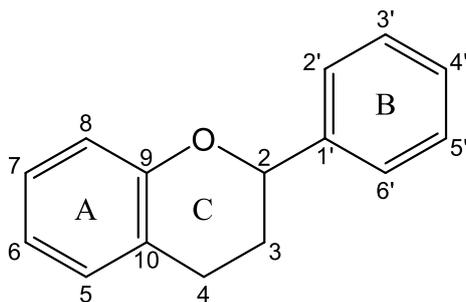


Figure 2 : Structure générale des composés flavonoïdes

### a) Les flavones, les flavanones, les flavonols et les flavononols

#### (1) Les flavones

Les flavones sont caractérisées par la présence d'une double liaison entre les carbones 2 et 3, et la présence d'un groupement cétone sur le carbone 4 (Figure 3). Le principal composé dans les vins est la lutéoline (Figure 3), qui a été détectée dans des vins chinois de cépage Cabernet sauvignon, Cabernet franc, Marselan et Merlot avec une concentration maximale de 4 mg/L pour ce dernier (83).

Il existe aussi des isomères des flavones, appelés isoflavones, qui se différencient par la position du cycle aromatique B lié au carbone 3. Une isoflavone a été détectée dans les vins et les baies, la génistéine (84) (Figure 3). Ces composés sont retrouvés à faibles concentrations dans les vins.

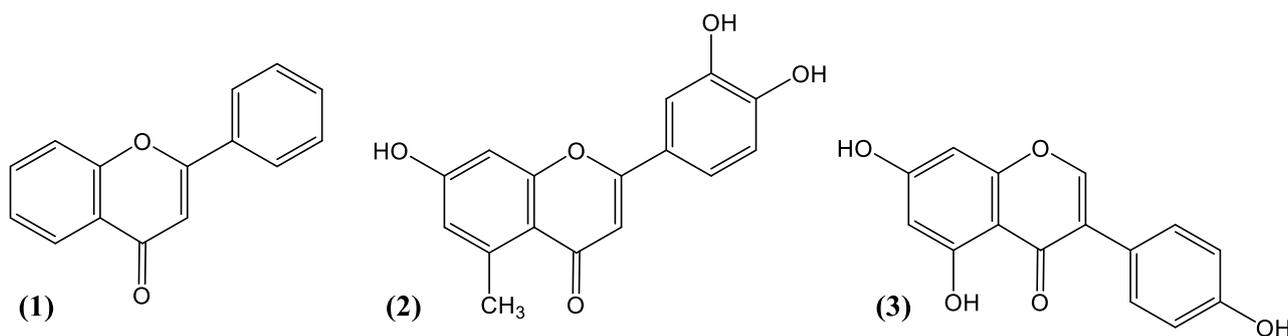


Figure 3 : (1) structure générale des flavones, (2) lutéoline, (3) génistéine

#### (2) Les flavanones

Les flavanones reprennent la structure principale des flavones, mais l'insaturation entre les carbones 2 et 3 disparaît. De ce fait, par analogie avec les flavones, ils sont souvent appelés dihydroflavones (85). Deux de ces composés ont été identifiés dans des vins tchèques, la naringinine

et l'hespérétine, à des concentrations respectives maximum de 0,8 et 0,7 mg/L (86). Sur les quatre vins analysés de l'étude, seulement deux vins ont montré des concentrations assez élevées pour être détectées, ces composés n'apparaissant qu'à l'état de trace.

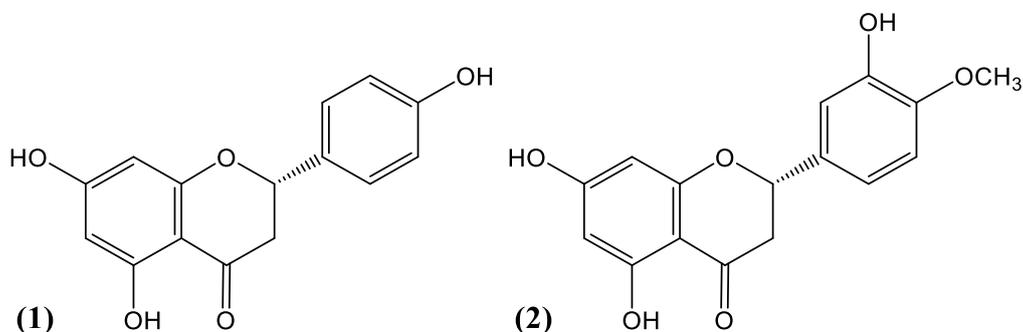


Figure 4 : Structures des flavanones retrouvés dans les vins : (1) naringine, (2) hespérétine

### (3) Les flavonols

Les flavonols reprennent la structure des flavones à laquelle vient s'ajouter un groupement hydroxyle sur le carbone 3. Il s'agit de pigments jaunes principalement localisés dans les vacuoles de la pellicule de la baie de raisin ainsi que dans les feuilles, avec comme principal rôle la protection face aux radiations UV (48,52). On peut également les retrouver en très faibles quantités dans la pulpe (87), mais aucune étude ne fait état de leur présence dans le pépin (88). Dans les grappes de *Vitis vinifera* sont retrouvés quatre principaux flavanols, la myricétine, la quercétine, le kaempférol, l'isorhamnétine, ainsi que leur forme glycosylée (89).

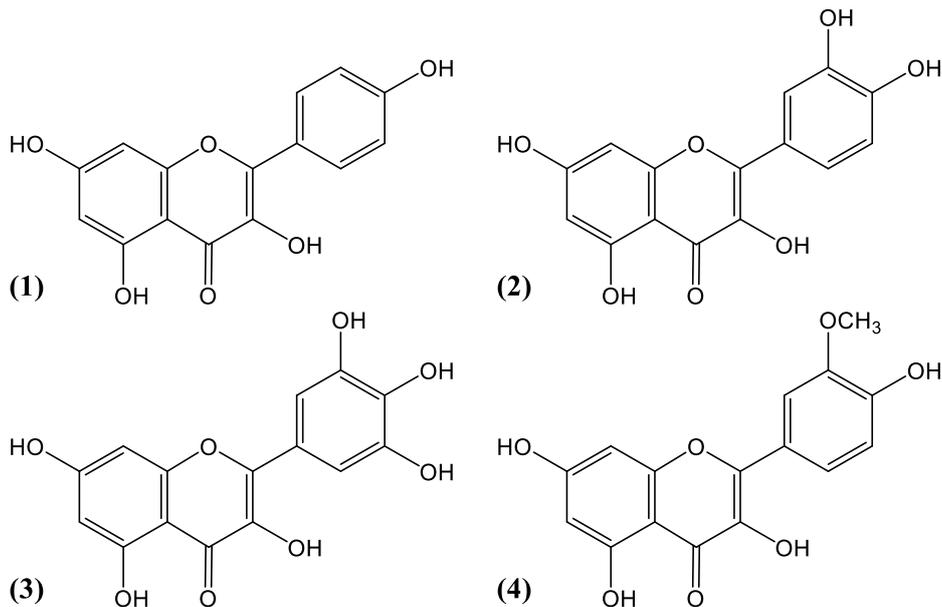


Figure 5 : Structures des principaux flavonols aglycones : (1) kaempférol, (2) quercétine, (3) myricétine, (4) isorhamnétine

Les flavonols peuvent être glycosylés par différents sucres sur le groupement hydroxyle porté par le carbone 3, les groupements glucosides, glucuronides et galactosides étant majoritaires (84). Si les concentrations en flavonols des cépages noirs peuvent varier, avec des teneurs minimum inférieures à 4 mg/kg et des teneurs maximum autour de 80 mg/kg, les composés majoritaires restent généralement les mêmes, avec près de 70 % de la concentration totale due à la myricétine et la quercétine (90). De même, il a été montré que dans les vins rouges les formes aglycones représentent deux tiers de la teneur totale en flavonols, avec dans le tiers restant les formes glycosylées majoritairement par du glucose (91).

#### (4) Les flavanonols

Les flavanonols sont caractérisés par l'absence de double liaison sur l'hétérocycle de la structure flavonoïde, et par un groupement hydroxyle sur le carbone 3. Ils sont aussi nommés 3-hydroxyflavonones ou dihydroflavonols. Le principal composé représenté dans les vins est la taxifoline ou dihydroquercétine (Figure 6), présents en très faible quantité avec des concentrations inférieurs à 1,5 mg/L pour des vins de Merlot ou de Cabernet franc (84,92,93).

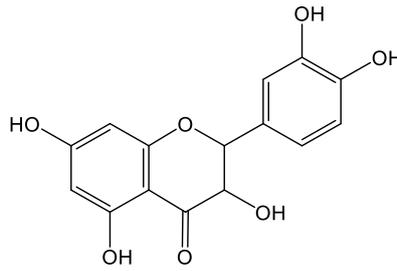


Figure 6 : Structure de la taxifoline

## ***b) Les flavan-3-ols et tanins condensés***

### *(1) Les monomères de flavan-3-ols*

Les flavanols, appelés plus précisément flavan-3-ols en raison du groupement hydroxyle en position 3 de l'hétérocycle, sont les composés flavonoïdes majoritaires du raisin et du vin, avec des concentrations comprises généralement entre 2 et 11 g/kg dans les baies de cépages noirs, et entre 1 et 4 g/L dans les vins rouges, les teneurs totales excédant rarement 50 % des teneurs de la baie (5,93,94). Dans le raisin, ces concentrations sont très variables et sont influencées par la nature du cépage (83,95), le climat ou le terroir (96,97).

Les principaux monomères des flavan-3-ols retrouvés dans le vin sont la (+)-catéchine, son isomère l'(-)-épicatéchine, l'(-)-épigallocatechine, l'(-)-épicatéchine-3-*O*-gallate et dans une moindre mesure la (+)-gallocatechine (98–100)(Figure 7). Il a aussi été identifié à des concentrations très inférieures aux monomères cités précédemment la (+)-gallocatechine-3-*O*-gallate et la (+)-catéchine-3-*O*-gallate (101). Cependant, ces composés n'ont été identifiés qu'à des concentrations très faibles, et joueraient donc un rôle mineur en comparaison des cinq flavan-3-ols principaux. Plus récemment, la présence d'épicatéchine-3-*O*-vanillate dans les raisins et vins a aussi été démontrée (102). Si ce composé se retrouve à des teneurs très faibles, inférieures à 250 µg/L, sa mise en évidence démontre la diversité des flavan-3-ols retrouvés dans le raisin et le vin.

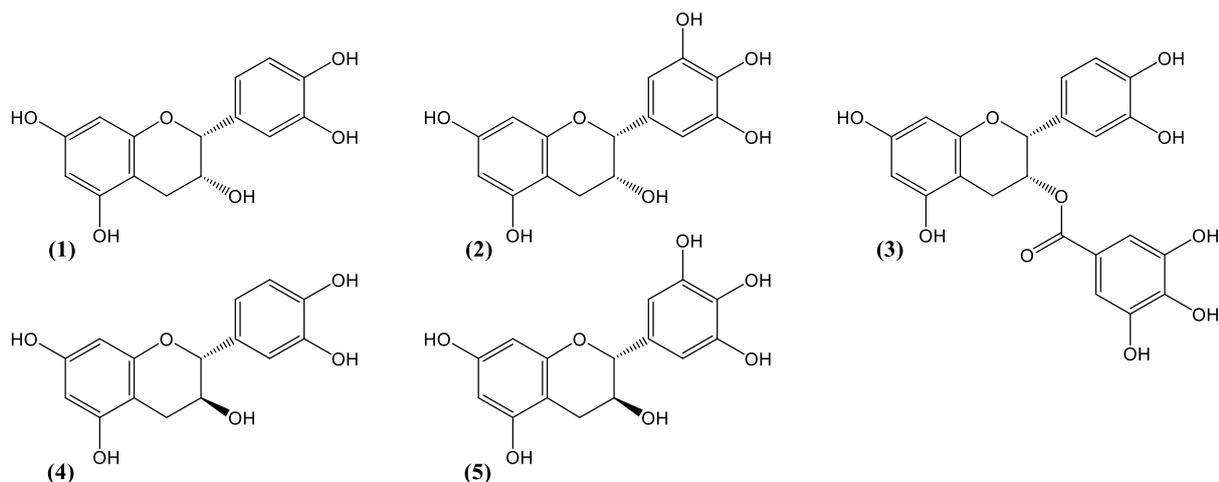


Figure 7 : Structures des principaux Flavan-3-ols du vin : (1) (-)-épicatéchine, (2) (-)-épigallocatechine, (3) (-)-épicatéchine-3-O-gallate, (4) (+)-catéchine, (5) (+)-gallocatechine

Les monomères de flavan-3-ols peuvent être retrouvés sous leur forme libre dans le raisin, bien que leurs formes polymérisées soient extrêmement prédominantes (97). Une étude menée sur du Cabernet franc a montré des teneurs en formes polymérisées près de vingt fois supérieures à celles en monomères (103). L'accumulation de ces composés ne se fait que dans des parties précises de la vigne, à savoir le pépin, la pellicule et la rafle. Dans la pellicule, ils sont essentiellement retrouvés dans les couches hypodermiques des cellules intérieures (104), sous forme libre, ou complexés avec les protéines des membranes vacuolaires ou les polysaccharides des parois cellulaires (105). Dans le pépin, ils sont situés dans le parenchyme externe, au niveau des parois cellulaires (106,107). De manière générale dans la baie de raisin, la plus grande proportion de flavan-3-ol se situe dans le pépin, à hauteur de 89 %, les 11 % restant se retrouvant dans la pellicule (108).

La structure des monomères est elle aussi influencée par la localisation dans la baie. Ainsi, l'épigallocatechine n'est présente que dans la pellicule, et n'a pas été retrouvée dans le pépin (109,110). A contrario, l'épicatéchine-3-O-gallate est un marqueur des pépins de raisin, sa concentration est largement supérieure à celle retrouvée dans la pellicule (111). Sa concentration peut ainsi être utilisée afin d'évaluer l'extraction globale des tanins de pépins lors de la vinification (109).

## (2) Les tanins condensés ou proanthocyanidines

Les oligomères et polymères des flavan-3-ols sont communément appelés tanins condensés ou proanthocyanidines. Ici, la définition d'oligomère et polymère est spécifique à la chimie du vin. Les oligomères sont formés par les molécules avec un degré de polymérisation compris entre 2 et 5, et les polymères sont composés par toutes les molécules plus importantes (52). Le terme

« tanins », quant à lui, réfèrent aux propriétés de ces derniers à donner des complexes stables avec les protéines, les polysaccharides ou encore les alcaloïdes (112,113). Pour finir, le terme proanthocyanidine indique la faculté des tanins condensés à se transformer en anthocyanidines en milieu acide et sous l'effet de la chaleur (114).

En fonction des unités flavan-3-ols qui les composent, les tanins condensés du vin peuvent être divisés en deux sous-groupes, les procyanidines et les prodelphinidines. Tout comme le terme proanthocyanidine, ces deux appellations trouvent leur origine dans la nature du composé formé en milieu acide et sous l'effet de la chaleur. Ainsi, les procyanidines sont composées de monomères catéchine et épicatechine, alors que les prodelphinidines sont composées d'unités d'épigallocatechine et gallocatechine (48).

Les tanins condensés se distinguent aussi par le type de liaison interflavane qui les composent. Deux types de liaisons ont été identifiés pour les proanthocyanidines, les liaisons de type B et les liaisons de type A, ces dernières étant fortement minoritaires (115)(Figure 8). Les liaisons de type B se composent de deux variantes : la liaison C4-C8, qui relie le carbone 4 de l'unité supérieure au carbone 8 de l'unité inférieure, et la liaison C4-C6, bien moins abondante que la première (116). Les liaisons de type A sont identifiables par l'une des liaisons C-C précédentes, à laquelle vient s'ajouter un pont éther entre le carbone 2 de l'unité supérieure et le groupement hydroxyle du carbone 5 ou 7 (52,117). Les deux ponts éther sont possibles indépendamment de la liaison de type B, ce qui donne au total pour les tanins condensés six liaisons différentes, deux de type B majoritaires et quatre de type A minoritaires.

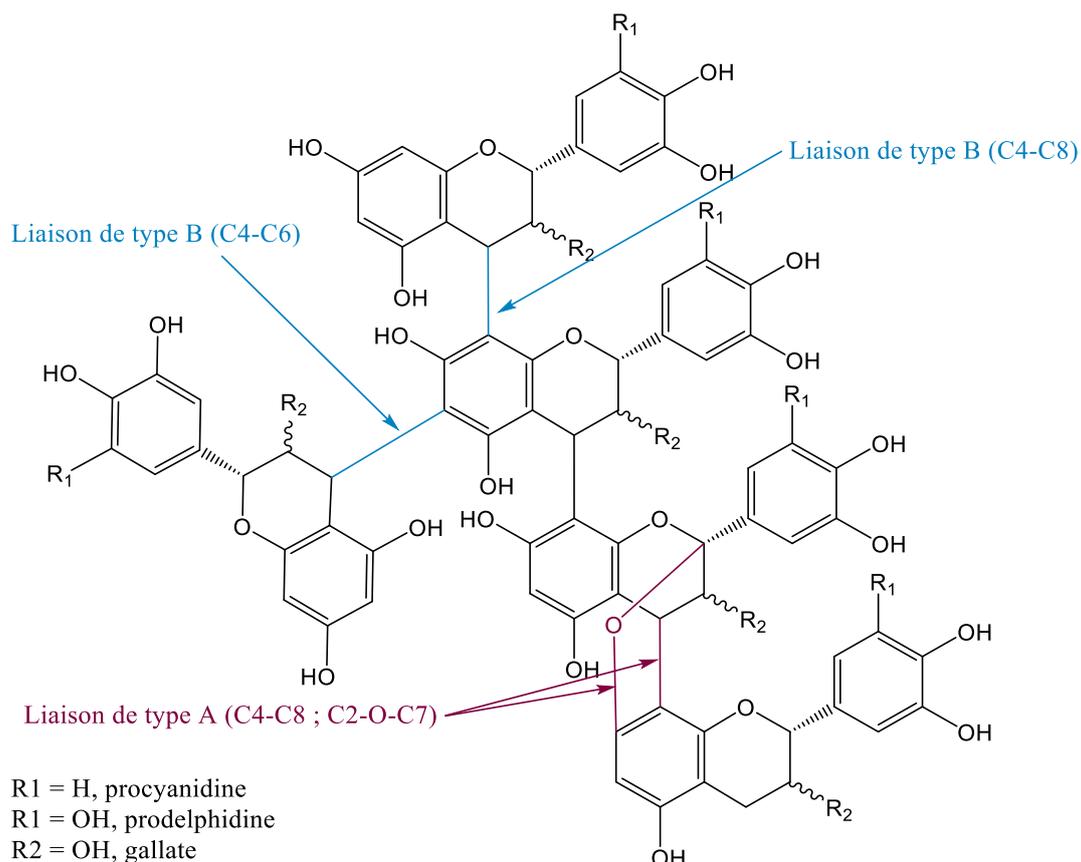


Figure 8 : Différents types de liaisons composants les tanins condensés

Un nombre extrêmement important de dimères et trimères ont été identifiés dans la baie, qu'il s'agisse du pépin ou de la pellicule, ainsi que dans le vin. Bien que les oligomères de procyanidines soient prédominants, c'est au total plus d'une quarantaine de composés différents qui ont été caractérisés au fil des études, mettant en lumière toute la diversité des tanins condensés du vin (48,52,84).

Les tanins du pépin et de la pellicule de raisin diffèrent dans leur composition en unités flavan-3-ols, mais aussi dans leur degré de polymérisation moyen (118). Ainsi, le degré de polymérisation moyen des tanins de la pellicule est plus important que celui des pépins, avec une moyenne des chaînes polymériques autour des 30 unités, contre 10 pour le pépin (119). Lors de la vinification, les grandes chaînes polymériques sont rarement extraites du matériel végétal, et les plus grandes identifiées dans les vins n'excèdent pas les 9 unités (120). De plus, il a été démontré récemment que le temps de macération n'a que peu d'impact sur le degré de polymérisation moyen des tanins condensés extraits des pépins ou de la pellicules (121). L'évolution de ce degré sur des macérations longues sera donc le résultat de l'évolution du ratio d'extraction entre pépin et pellicule.

### (3) Les tanins couronnes

Récemment une nouvelle classe de tanins condensés, les tanins couronnes, a été identifiée dans le vin et le raisin. Le premier composé de cette classe a été caractérisé par spectrométrie de masse et analyse par résonance magnétique nucléaire. Il s'agit d'un tétramère de procyanidine, et il se différencie des structures procyanidiques linéaires classiques par une structure cyclique (122) (Figure 9). Neuf autres proanthocyanidines cycliques ont été partiellement identifiées dans les vins, notamment des formes pentamériques, mais l'analyse par spectrométrie de masse seule nécessite une investigation plus poussée pour valider la structure de ces nouveaux composés (123).

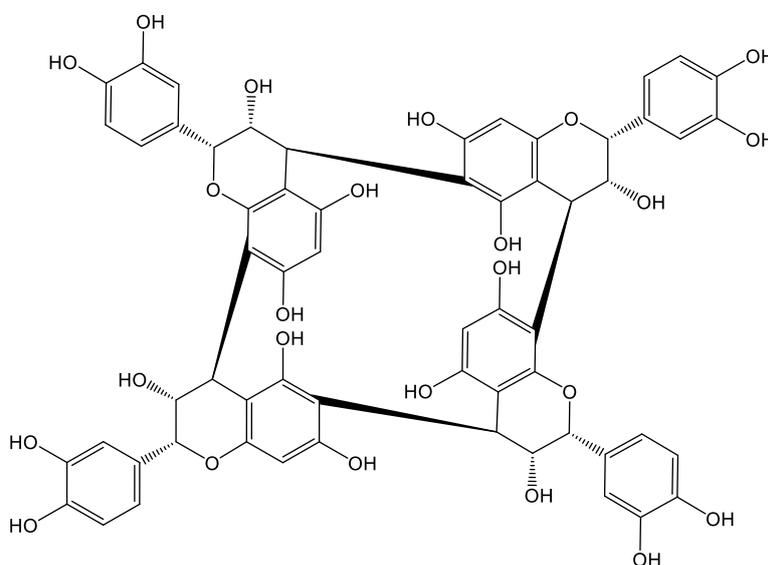


Figure 9 : Structure d'un tétramère de procyanidine cyclique

La concentration en tétramère couronne a été suivie sur une verticale de vin du Languedoc couvrant 20 millésimes et comparée à son homologue linéaire (124). Contrairement au tétramère cyclique, les tanins couronnes ne semblent pas sujet à la transformation et restent stables au fil des années, avec une concentration proche de 40 mg/L. D'autres paramètres pouvant influencer sur les tétramères couronnes ont aussi été investigués, notamment l'effet cépage sur les teneurs totales de ces composés et sur le ratio tétramère couronnes/tétramère linéaire (125,126). Les résultats mettent en évidence des différences de concentration en fonction du cépage, mais surtout une variabilité importante du ratio couronne/linéaire. Ce dernier apparaît comme un outil prometteur dans l'authentification des vins.

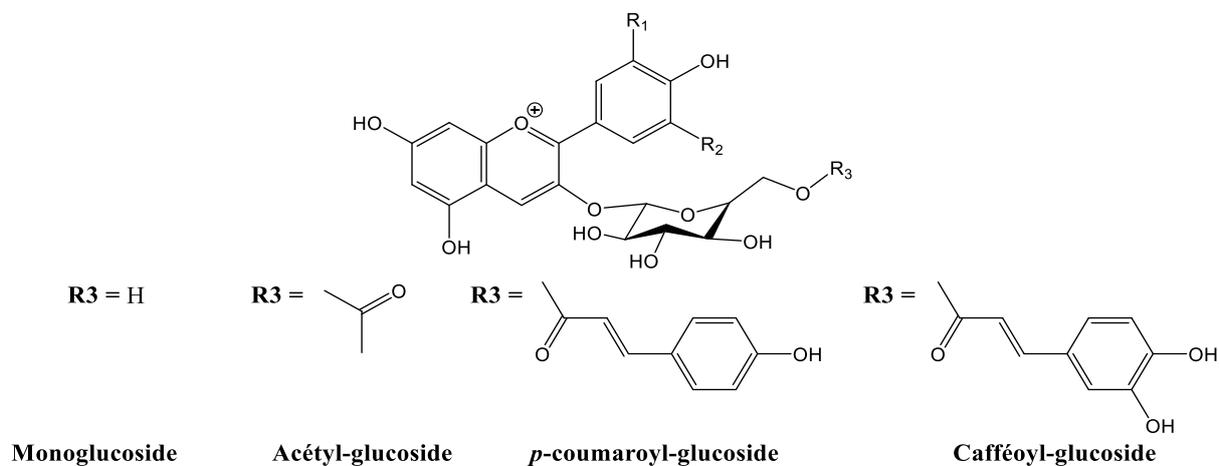
### c) Les anthocyanes

#### (1) Structures, localisations et concentrations

Le terme anthocyane, du grec *anthos* pour fleur et *kianos* pour bleu, a été proposé pour la première fois par Macquart en 1835 pour désigner les composés responsables de la couleur bleue des bleuets (127). En réalité, les anthocyanes représentent la plus importante classe de pigments solubles du règne végétal, et sont responsables de pratiquement toutes les nuances de rouge, bleu et violet retrouvées dans les plantes. Après les tanins condensés, les anthocyanes sont la deuxième classe de polyphénol la plus importante en termes de concentration dans les vins et elles sont responsables de la couleur du raisin noir et du vin rouge. Elles se composent d'une anthocyanidine et d'un ou plusieurs sucres. Les formes majoritaires des vins et raisins sont monoglucosides, avec un sucre positionné sur l'hydroxyle du carbone 3 (4,128). Dans les vins issus de *Vitis vinifera*, cette concentration varie en général autour de 500 mg/L, avec des maximums pouvant dépasser les 1500 mg/L (129–131).

Dans le raisin, elles sont uniquement localisées dans la pellicule, plus précisément dans les vacuoles des cellules épidermiques (132). Certains cépages, dits « teinturiers », font exception et contiennent également des anthocyanes dans la pulpe de la baie (133,134).

Cinq anthocyanes principales ont été identifiées dans le raisin et le vin : la malvidine-3-*O*-glucoside, la cyanidine-3-*O*-glucoside, la péonidine-3-*O*-glucoside, la pétunidine-3-*O*-glucoside et la delphinidine-3-*O*-glucoside. Ces composés diffèrent les uns des autres par le nombre et la position des groupements hydroxyles et méthoxyles du cycle B. Une acylation peut également se produire en position 6 du glucose, par estérification avec de l'acide acétique, de l'acide *p*-coumarique ou moins communément avec de l'acide caféique (135).



<b>Anthocyane</b>	<b>R<sub>1</sub></b>	<b>R<sub>2</sub></b>
Cyanidine-3- <i>O</i> -glucoside	OH	H
Delphinidine-3- <i>O</i> -glucoside	OH	OH
Péonidine-3- <i>O</i> -glucoside	OCH <sub>3</sub>	H
Pétunidine-3- <i>O</i> -glucoside	OCH <sub>3</sub>	OH
Malvidine-3- <i>O</i> -glucoside	OCH <sub>3</sub>	OCH <sub>3</sub>

Tableau 3 : Structures des principales anthocyanes du raisin et du vin

D'autres dérivés des anthocyanes ont été identifiés dans les vins, tels que la pélargonidine-3-*O*-glucoside, des anthocyanes acétylées avec de l'acide lactique, ou encore des anthocyanes diglucosides (136–138). Ces composés sont cependant retrouvés à des concentrations faibles en comparaison des anthocyanes citées précédemment. La pélargonidine-3-*O*-glucoside, par exemple, a été détectée dans les vins à l'état de trace, et semble absente du jus de raisin (139,140). Une étude menée sur le cépage teinturier Alicante Bouschet a montré qu'elle représentait moins de 0,5 % du pourcentage molaire des anthocyanes totales (137).

Tout comme pour les tanins condensés, les distributions et concentrations des anthocyanes sont dépendantes du cépage, de la maturité, des conditions climatiques et du terroir (95,141,142). En général, la malvidine-3-*O*-glucoside est le composé majoritaire chez *Vitis vinifera*, composant près de 90 % des anthocyanes dans les vins de Grenache, mais moins de 50 % des vins de Sangiovese (48). Cependant, il a été démontré récemment que pour certaines espèces hybrides de vigne, l'anthocyane prédominante pouvait changer. C'est notamment le cas pour le cépage Regent, dans lequel la delphinidine-3-*O*-glucoside a été identifiée comme majoritaire (143). Les anthocyanes acétylées font elles aussi preuve d'une grande variabilité, et peuvent être absentes de certains cépages, comme pour le Pinot noir (144).

## (2) Equilibre des anthocyanes en fonction du pH

La couleur d'une solution en anthocyanes monomériques est dépendante de son pH. En solution acide, c'est la forme flavylium  $A^+$  qui est prédominante, et possède une couleur rouge caractéristique. En augmentant le pH, la stabilité de la forme flavylium est perturbée et se dirige vers la base carbinol AOH incolore, la chalcone C jaune ou la base quinonique AO bleu-mauve (128,145,146)(Figure 10). Au pH du vin, compris entre 3 et 4, les quatre formes co-existent. La forme incolore carbinol est alors prédominante, avec des proportions comprises entre 35 et 50 %. La chalcone se retrouve alors être le chromophore majoritaire et varie de 20 à 30 %. La base quinonique est la forme la plus stable au pH du vin, avec une variabilité réduite entre 15 et 20 %. Pour finir, la forme flavylium, contributrice majoritaire de la couleur rouge des vins, est celle qui subit les variations les plus importantes. A pH 3, elle compose près de 35 % de la proportion totale, pour passer en dessous de 10 % une fois le pH 4 dépassé.

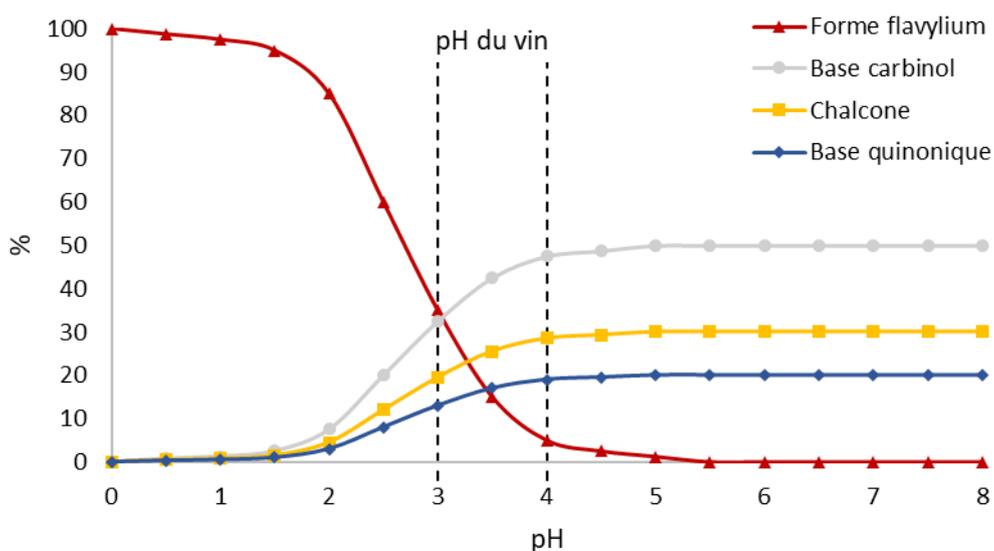


Figure 10 : Evolution des différentes formes des anthocyanes en fonction du pH

Ces équilibres sont induits par différentes réactions survenant dans un milieu de l'acidité du vin. Le cation flavylium est alors sujet à deux types de réactions : une réaction acide-base et une réaction d'hydratation (145). Il se comporte soit comme un électrophile via les carbones 2 et 4 du cycle B ou comme un nucléophile via les carbones 6 et 8 du cycle A (48). La réaction acide-base conduira à la formation possible de deux quinones distinctes, et la réaction d'hydratation à la formation de deux bases carbinols. C'est l'une de ces dernières qui pourra ensuite subir une ouverture de son hétérocycle

pour conduire à l'apparition de la chalcone, en équilibre isomérique entre la forme *cis*- et *trans*- (Figure 11).

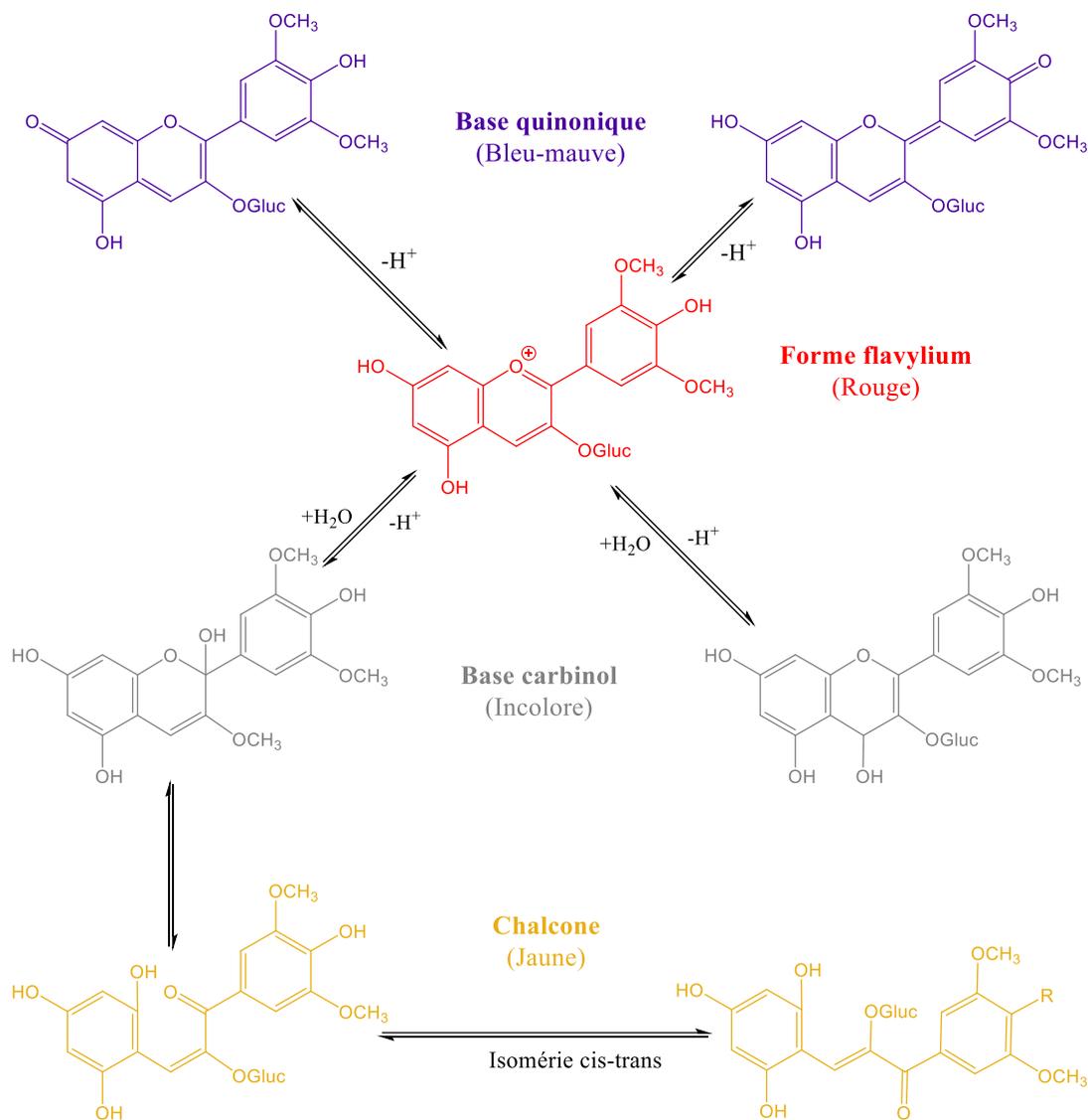
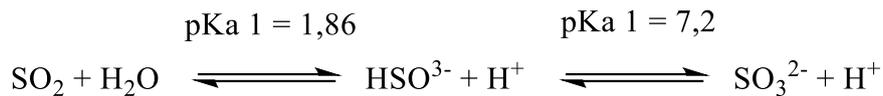


Figure 11 : Equilibre entre les différentes formes anthocyaniques : l'ion flavylium, la base quinonique, la base carbinol et la chalcone, exemple de la malvidine-3-*O*-glucoside

### (3) Décoloration des anthocyanes en présence de bisulfite de sodium

Le bisulfite de sodium ( $SO_2$ ) est un produit quasiment indispensable à l'œnologie, grâce à son action directe et préventive sur l'activité antimicrobienne, et protectrice contre l'oxydation (147,148). En milieu aqueux, le  $SO_2$  peut réagir avec l'eau en formant le couple acido-basique  $SO_2$ ,  $H_2O/HSO_3^-$ , puis le couple  $HSO_3^-/SO_3^{2-}$  et conduit à des équilibres avec les constantes  $pK_a$  1 et  $pK_a$  2 correspondantes :



Au pH du vin, le  $\text{SO}_2$  ne se retrouve pratiquement que sous la forme de l'ion hydrogénosulfite  $\text{HSO}_3^-$  où il représente 96,3 % de la proportion totale à pH 3,2 et 98 % à pH 4 (149). Sous cette forme, il possède un fort pouvoir décolorant sur les anthocyanes monomériques (150). Ainsi, l'ion hydrogénosulfite peut réagir avec la forme flavylum des anthocyanes sur le carbone 2 ou 4, pour former deux composés incolores possibles (151)(Figure 12).

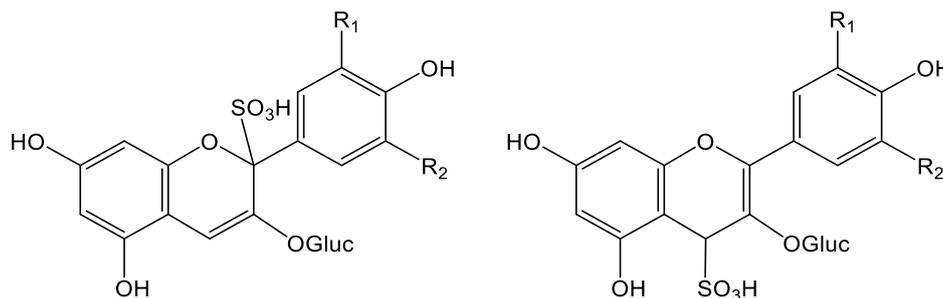


Figure 12 : Structures des produits de la réaction entre l'ion hydrogénosulfite et les anthocyanes

### 3. Les tanins hydrolysables

#### a) Classification et structures des tanins hydrolysables

Au cours de l'élevage des vins en barriques, des tanins présents dans le bois peuvent être extraits et solubilisés. A l'inverse des tanins condensés de la baie de raisin, ces tanins dit hydrolysables, par leur faculté à être hydrolysés peuvent libérer soit de l'acide gallique pour les gallotanins soit de l'acide ellagique pour les ellagitanins, qui ne sont pas présents naturellement dans le raisin (152,153). Le précurseur de ces tanins est le 1,2,3,4,6-penta-*O*-galloyl- $\beta$ -D-glucopyranose (154), et les gallotanins sont définis par ce noyau sur lequel sont estérifiées une à plusieurs fonctions galloyles, formant un à plusieurs groupements digalloyl depsidiques caractéristiques (155) (Figure 13). Les gallotanins les plus complexes sont le résultat de nombreuses réactions de galloylation et les métabolites formées peuvent contenir jusqu'à plus de dix résidus galloyles (156). Si les gallotanins ne se retrouvent pas ou peu dans le bois de chêne, principale source pour l'élevage des vins en barriques, ils sont couramment utilisés comme tanins œnologiques. Ils ont ainsi démontrés un caractère très bénéfique quand ils sont ajoutés lors de la vinification : ils permettent de stabiliser la couleur par copigmentation (157,158), ils ont une forte capacité antioxydante (159) et ont une action inhibitrice de l'activité laccase qui peut prévenir des dégâts liés à *Botrytis cinerea* (160,161).

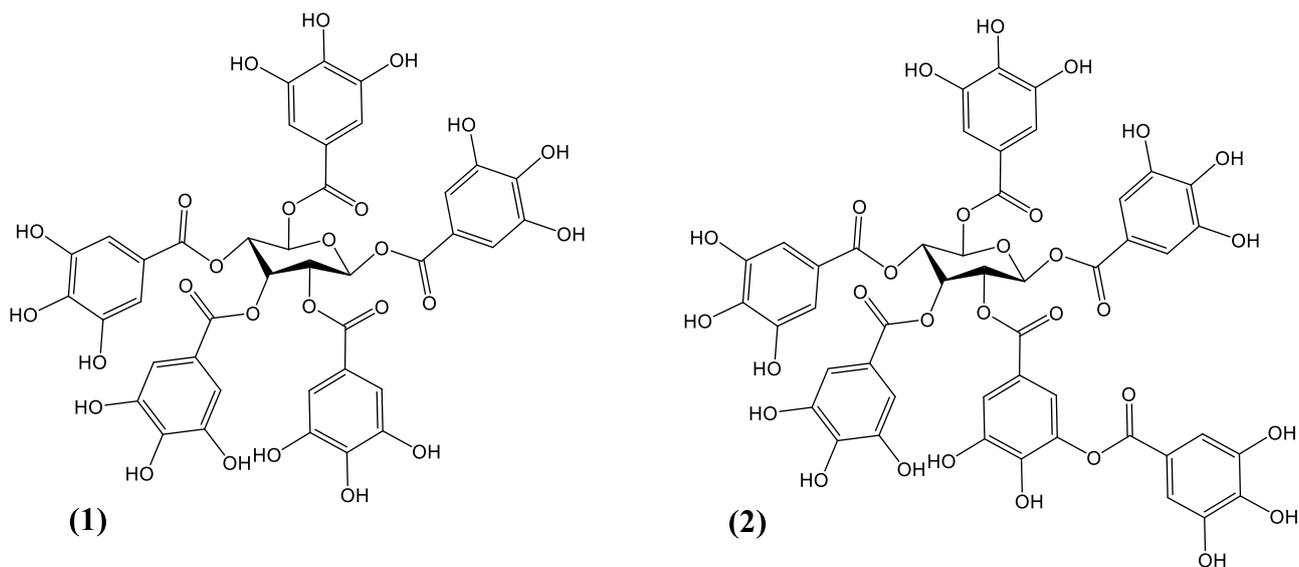


Figure 13 : (1) 1,2,3,4,6-penta-*O*-galloyl- $\beta$ -D-glucopyranose, précurseur des tanins hydrolysables ; (2) exemple de gallotanins, le 2-*O*-digalloyl-1,2,3,4,6-tétra-*O*-galloyl- $\beta$ -D-glucopyranose

Le noyau de 1,2,3,4,6-penta-*O*-galloyl- $\beta$ -D-glucopyranose peut aussi être sujet à des oxydations. Lorsque cela se produit, des liaisons biaryl se forment entre groupements galloyls pour conduire à un groupement 3,4,5,3',4',5'-hexahydroxydiphénol (HHDP), caractéristique de la seconde classe de tanins hydrolysables, les ellagitanins (Quideau & Feldman, 1996). Lorsqu'ils sont hydrolysés, ces composés libèrent de l'acide hexahydroxydiphénique qui se transforme spontanément dans la dilactone correspondante, l'acide ellagique, qui donne son nom aux ellagitanins (156).

Dans le bois de chêne, et en conséquence dans les vins élevés en barriques, on peut retrouver huit ellagitanins C-glycosidiques principaux : les monomères castalagine, vescalagine, grandinine et roburine E, et les dimères roburines A,B,C et D (162,163). Les ellagitanins C-glycosidiques sont une sous-classe d'ellagitanins qui présentent la particularité structurale d'avoir une liaison C-C caractéristique entre le carbone C-1 de la chaîne ouverte du noyau glucose et le carbone C-2 de l'unité 2-*O*-galloyl (154) (Figure 14).

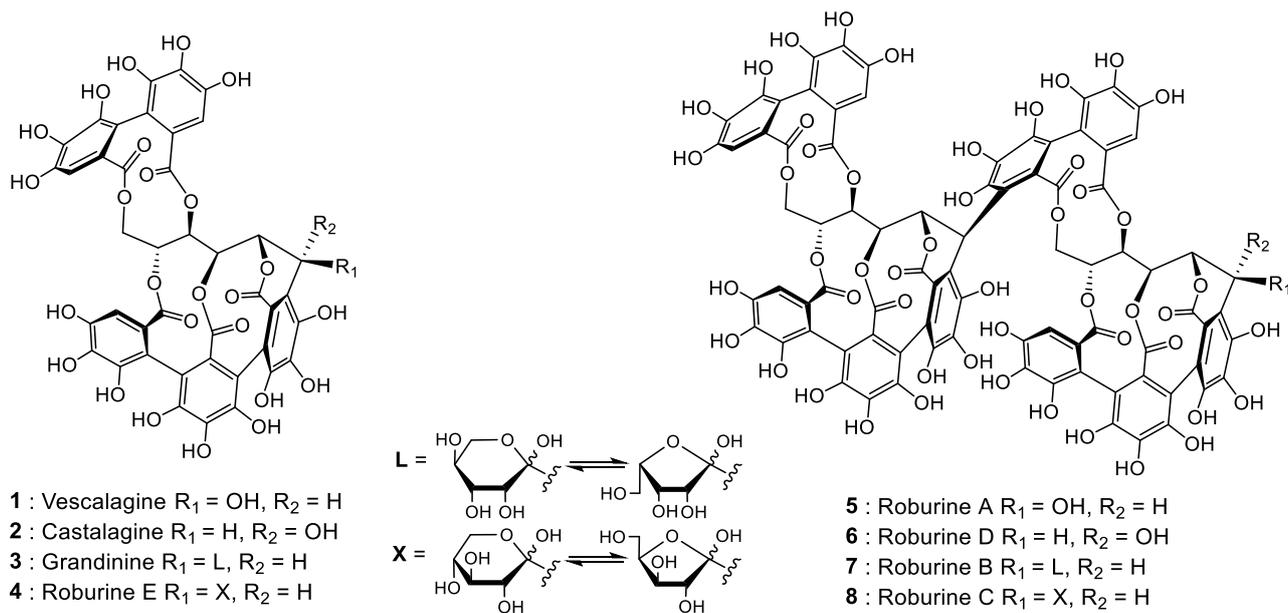


Figure 14 : Structures des 8 ellagitannins C-glycosidiques principaux présents dans le bois de chêne

### **b) Extraction et évolution des ellagitannins dans le vin**

L'élevage des vins rouges en présence de bois permet de moduler les propriétés organoleptiques du produit, en amenant de la complexité au niveau aromatique, en modifiant les perceptions d'astringence et d'amertume et en stabilisant la couleur (164–166). Cet élevage est généralement conduit en barriques de chêne pendant 6 à 18 mois, mais ce processus destiné aux vins qualitatifs est coûteux et il peut lui être préféré un élevage en cuve dans lesquelles sont ajoutés des morceaux de bois de chêne (copeaux, staves, ...). Le vin, en contact avec le bois, va ainsi pouvoir extraire et solubiliser les ellagitannins C-glycosidiques présents dans le chêne. L'extraction est influencée par différents facteurs. Ainsi, l'espèce de chêne utilisée va impacter grandement les teneurs finales en ellagitannins (167). Sur les trois grandes familles d'espèces de chênes, les espèces européennes *Quercus robur* et *Quercus petraea* démontrent des niveaux en ellagitannins dans les vins bien supérieurs à *Quercus alba*, l'espèce américaine (168,169). Outre l'espèce de chêne, plusieurs paramètres ont aussi une influence sur les teneurs en ellagitannins dans le bois : la provenance géographique de l'arbre, la chauffe du bois, la position de l'échantillon dans l'arbre ou encore les méthodes de séchage des staves (167,170–172). Les barriques étant souvent utilisées pendant plusieurs campagnes de vinification, le nombre d'utilisation est aussi à prendre en compte car le niveau d'ellagitannins baisse drastiquement après la première vinification (169).

Au cours de l'élevage en barriques, la concentration en ellagitanins va augmenter jusqu'à un maximum pour ensuite diminuer lentement (Figure 15). Lors des premiers mois, l'extraction est importante puis se retrouve contrebalancée par les réactions de dégradation ou de complexation avec d'autres composés du vin, notamment les flavonoïdes (163,173). Due à ces transformations, la teneur en ellagitanins libres d'un vin fini est faible. Ainsi, si les concentrations peuvent atteindre un maximum de 50 mg/L pendant la vinification, après seulement 12 mois cette concentration est réduite de plus de la moitié (Figure 15). Lors du vieillissement en bouteilles, cette évolution continue pour atteindre des teneurs inférieures à 5 mg/L après 4 ans et inférieur à 1 mg/L après 12 ans (174).

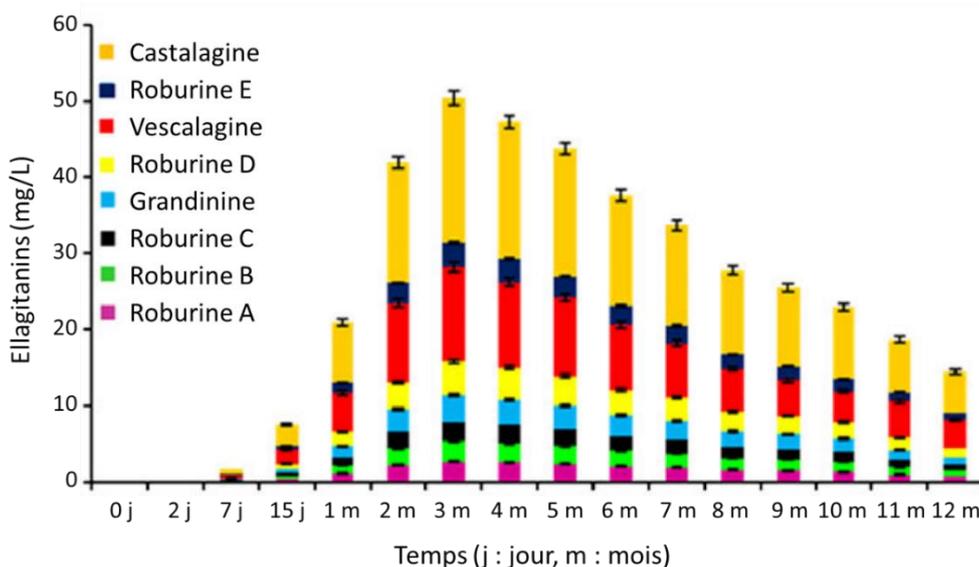


Figure 15 : Evolution de la teneur en ellagitanins C-glycosidiques durant l'élevage en barriques de chêne (163)

### c) Réaction des ellagitanins avec les flavonoïdes

Les ellagitanins, notamment les composés avec un encombrement stérique peu important comme la vescalagine, sont sensibles à une déshydroxylation de l'alcool benzilique situé sur le carbone 1 du glucose ouvert. Cette réaction est favorisée en milieu acide, comme celui du vin, et conduit à la formation d'un carbocation hautement réactif avec les composés nucléophiles (152). Ce carbocation peut être attaqué par les nombreuses molécules nucléophiles du vin, comme l'éthanol pour former avec la vescalagine de l'éthylvescalagine, mais aussi par les composés flavonoïdes : flavan-3-ol, tanins condensés et anthocyanes.

L'addition de flavan-3-ols conduit à la formation de flavano-ellagitanins, dont les 4 principaux retrouvés sont l'acutissimine A et B, fruits de l'addition de la vescalagine et la catéchine, et

l'épiacutissimine A et B, fruits de l'addition de la vescalagine et l'épicatéchine (175)(Figure 16). Le flavano-ellagitanin retrouvé en plus grande concentration est l'acutissimine A, mais les teneurs dans les vins restent très faibles, ne dépassant pas les 0,5 mg/L pour des vins élevés minimum 12 mois en barriques (174,175).

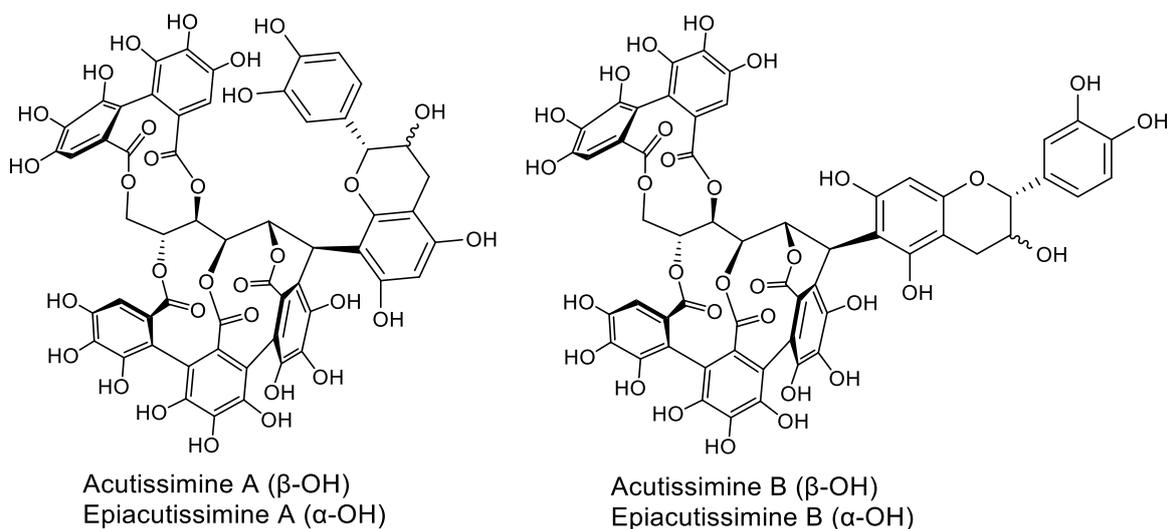


Figure 16 : Structure des principaux flavano-ellagitanins du vin

La vescalagine peut aussi réagir avec les anthocyanes, pour former des anthocyano-ellagitanins (152)(Figure 17). La principale observation qui peut être faite sur l'apparition de ces nouveaux composés est une différence notable du maximum d'absorption dans le visible, et donc de la couleur du milieu (176). A pH 1, la malvidine-3-*O*-glucoside montre un maximum d'absorption à 519 nm, et la malvidine-3-*O*-glucoside-vescalagine un maximum d'absorption à 542 nm, indiquant un effet bathochrome de la réaction d'addition. Cela se traduit par le passage d'un milieu rouge lorsque la malvidine-3-*O*-glucoside est libre à un milieu violet lorsque celle-ci réagit avec la vescalagine. S'il est impossible d'affirmer que le complexe anthocyane-ellagitanin est un acteur important de la couleur des vins, il semble cependant contribuer à sa modulation lors de l'élevage en présence de bois, et pourrait expliquer l'augmentation de la composante pourpre observée sur les vins riches en ellagitanins (152,166).

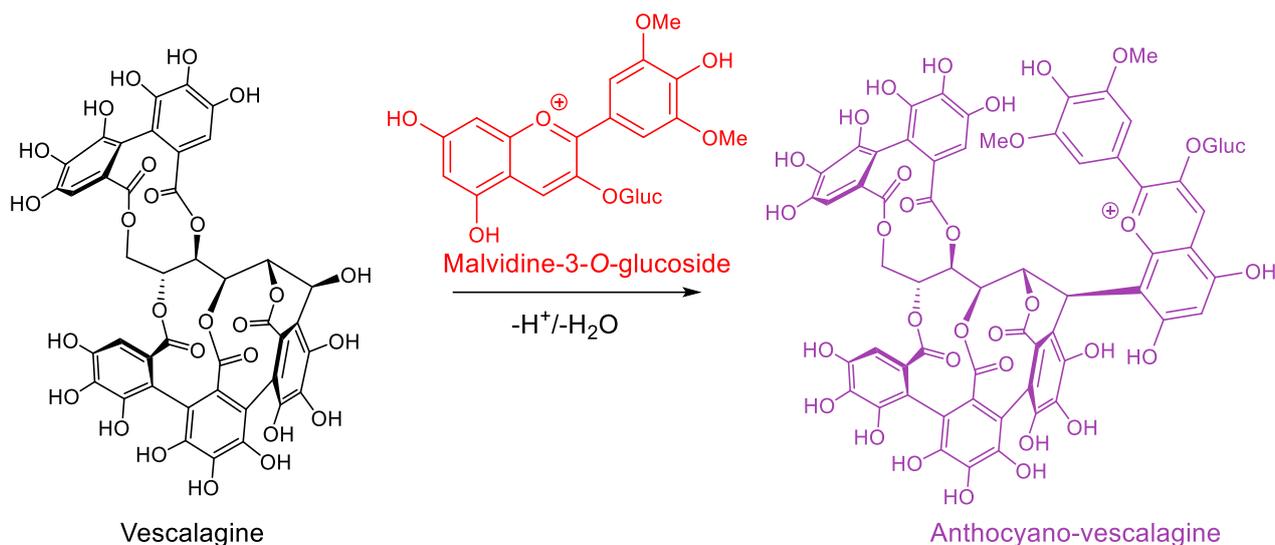


Figure 17 : Exemple de formation d'un anthocyano-ellagitanin, réaction de la vescalagine avec la malvidine-3-O-glucoside

#### 4. Evolution des flavonoïdes du vin : condensation et dégradation

Une fois extraits de la baie de raisin et solubilisés pendant la fermentation alcoolique, les composés flavonoïdes du vin sont amenés à évoluer au cours de la vinification et du vieillissement. Les tanins condensés et les anthocyanes, par leur réactivité, vont ainsi pouvoir subir des réactions de dégradation ou de polymérisation, fortement influencées par la variabilité de la matrice chimique du vin.

##### a) Réactions directes de condensation

###### (1) Polymérisation directe par catalyse acide

Des réactions de réarrangement peuvent se produire entre proanthocyanidines, par dépolymérisation et polymérisation des unités flavan-3-ols. Le clivage d'une liaison interflavane peut être induit par catalyse acide, et produit un intermédiaire carbocation (177). Ce carbocation très électrophile va pouvoir réagir avec tous composés nucléophiles du vin, en général une unité flavan-3-ol ou une anthocyane sous sa forme carbinol, et plus rarement avec des flavonols ou des thiols (178,179)(Figure 18).

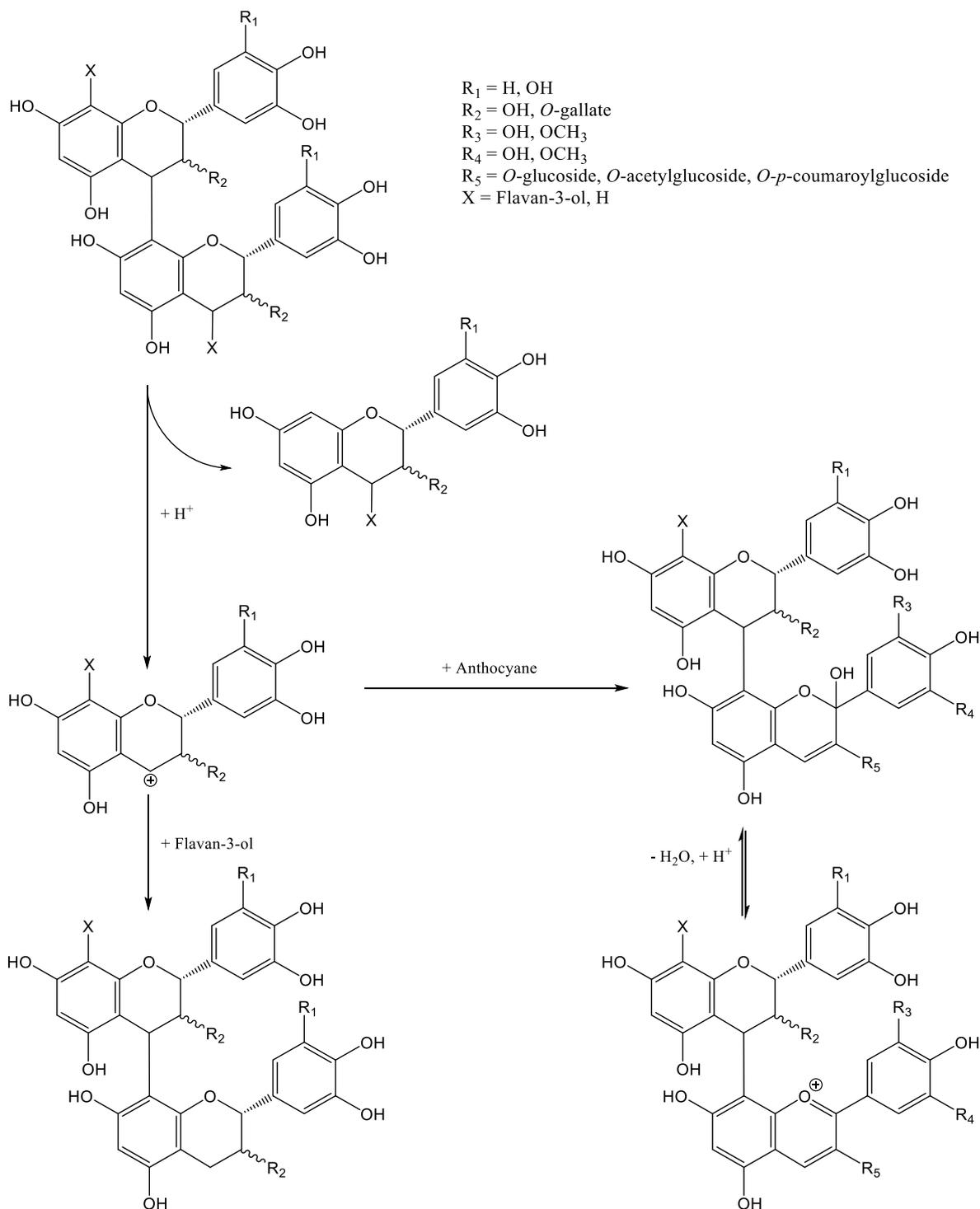


Figure 18 : Polymérisation directe par catalyse acide : réaction flavan-3-ol/flavan-3-ol et flavan-3-ol/anthocyane

Lorsque cette réaction a lieu en présence d'un excès de monomères de flavan-3-ol, le degré de polymérisation moyen va graduellement diminuer, les polymères laissant place à des oligomères (179). Quant à la réaction avec une anthocyane, elle produit dans un premier temps un intermédiaire

hémiacétal incolore qui après protonation et déshydratation va se réarranger en adduit flavylum, de couleur rouge-orangée (180).

### (2) Polymérisation directe en présence d'oxygène

En présence d'oxygène, les flavan-3-ols peuvent s'oxyder en quinone ou semi-quinone, espèces chimiques très réactives et instables. Lorsque cela survient sur un monomère ou une unité terminale de proanthocyanidine, une réaction de condensation est alors possible (181). Cette polymérisation se distingue des autres car la réactivité est délocalisée sur le cycle B et non sur l'hétérocycle C. Dans le cas où l'oxydation se fait sur une unité autre que terminale, cela conduit à la formation d'une liaison de type A (182)(Figure 19). Cette réaction est également possible avec une anthocyane en unité terminale.

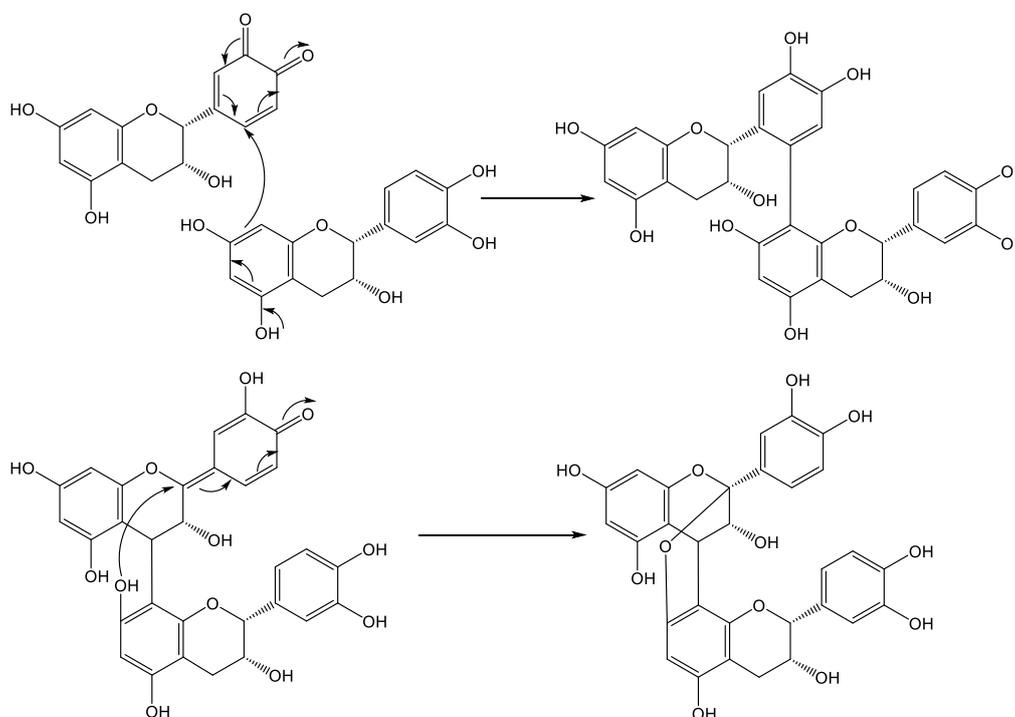


Figure 19 : Polymérisation directe et formation de liaison de type A, exemple de monomères et dimères d'épicatéchine

### (3) Polymérisation directe entre les flavan-3-ols et les anthocyanes

Lors de la vinification, les flavan-3-ols et les anthocyanes sont extraits et coexistent en solution. Des réactions de polymérisation surviennent alors naturellement entre ces composés, avec pour effet une évolution de la couleur du rouge vers des teintes tuilées. L'attaque nucléophile des

anthocyanes a déjà été évoquée lors de la polymérisation directe par catalyse acide, mais ces composés ont également des propriétés électrophiles sous leur forme de cation flavylum (183,184). De cette réactivité résulte la formation d'un composé anthocyane-flavan-3-ol, où l'anthocyane constituera la première unité du polymère, contrairement à l'attaque nucléophile où elle se place en unité terminale. Dans les deux cas, l'addition d'une anthocyane sur une proanthocyanidine bloque la possibilité de condensation ultérieure, celle-ci une fois liée présentant une réactivité extrêmement faible (185). La réaction a lieu entre le carbone 6 ou 8 du flavan-3-ol et le carbone 4 de l'anthocyane (186)(Figure 20).

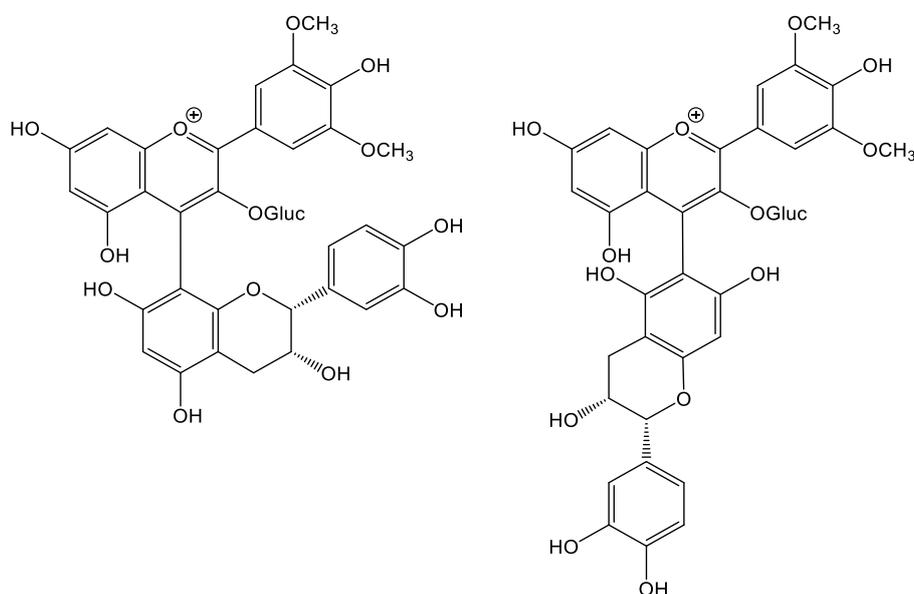


Figure 20 : Exemple de condensation directe entre la malvidine-3-O-glucoside et l'épicatéchine

En présence d'oxygène, ces composés peuvent également former une liaison de type A ou conduire à la formation de xanthyliums (187–189).

### ***b) Réactions indirectes de condensation en présence d'acétaldéhyde***

Différents aldéhydes (acétaldéhyde, acide glyoxylique, furfural, ...) sont rencontrés dans le vin. Le plus abondant est l'acétaldéhyde qui est à la fois un produit du métabolisme de la levure et un produit d'oxydation de l'éthanol (190–192). Ce composé peut réagir avec les flavonoïdes et induire des condensations indirectes par pont éthylidène. Ce mécanisme de réaction a été proposé pour la première fois par Timberlake en 1976 et validé vingt ans plus tard par spectrométrie de masse (193,194).

La protonation de l'acétaldéhyde conduit à un intermédiaire carbocation qui peut réagir sur le carbone 6 ou 8 d'un flavan-3-ol monomère ou condensé. Le composé flavano-acétaldéhyde obtenu peut alors subir une protonation et une déshydratation pour former un nouveau carbocation. Cette nouvelle molécule peut réagir avec les flavonoïdes ou autres nucléophiles présents dans le milieu et les lier par pont éthylidène (Figure 21). Sachant que les carbones 6 et 8 sont tous les deux réactifs, la réaction peut continuer sur le site nucléophile vacant et conduire à la polymérisation (195).

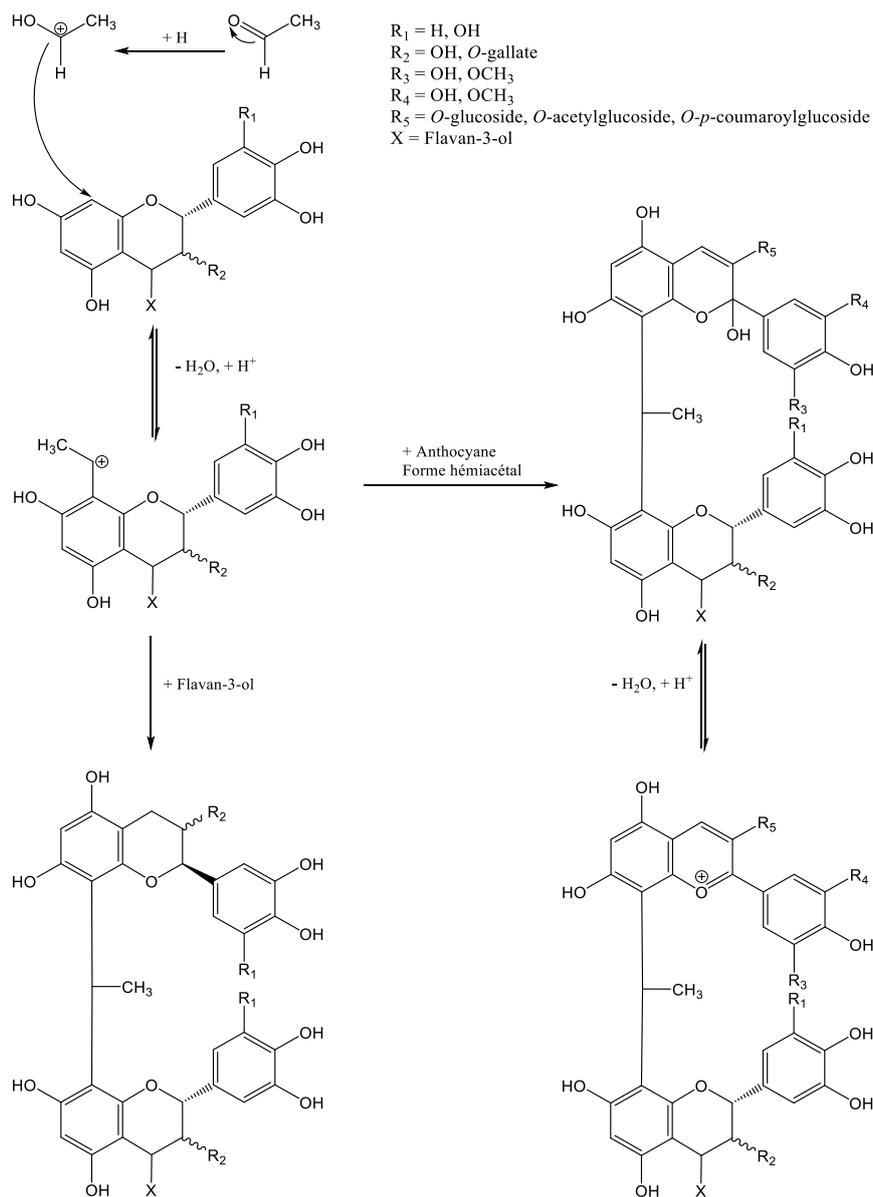


Figure 21 : Polymérisation indirecte en présence d'acétaldéhyde : réaction flavan-3-ol/flavan-3-ol et flavan-3-ol/anthocyane

Lorsque des flavan-3-ols et des anthocyanes sont présents dans le milieu, l'acétaldéhyde attaque préférentiellement les flavan-3-ols, ce qui a longtemps fait penser que les anthocyanes ne présentaient

pas de réactivité. Cependant, la détection d'anthocyanes liées par pont éthylidène a mis fin à cette hypothèse, bien que ces composés soient en proportion faible dans les vins en comparaison aux flavan-3-ols liés par pont éthylidène (196). De plus, les anthocyanes impliqués dans les réactions avec l'acétaldéhyde sont sous la forme hémiacétale du carbinol, incolore, alors que le composé formé est quant à lui coloré, ce qui conduit un renforcement général de la couleur des vins.

Deux études concentrées sur des vins de Bordeaux ont montré que la proportion de ponts éthylidènes entre flavan-3-ols augmentait lors du vieillissement (197,198). Cette proportion reste cependant faible en comparaison des liaisons directes, et représente moins de 0,5 % des liaisons entre flavan-3-ols sur des vins de 2 ans, pour évoluer jusqu'à un maximum proche de 4 % sur des vins supérieurs à 20 ans.

### ***c) Evolution des anthocyanes dans les vins***

#### ***(1) Dégradation des anthocyanes***

Les anthocyanes libres du vin montrent une stabilité chimique très relative, et sont sujettes à toutes sortes d'attaques nucléophiles et électrophiles par les différentes molécules réactives en solution (199). Ainsi, la concentration de ces composés va diminuer très fortement dès les premiers mois de la vinification. Si la grande majorité est impliquée dans les réactions de polymérisation avec les tanins condensés, elles peuvent aussi subir des réactions de dégradation. Deux dégradations possibles ont été identifiées dans les vins, la dégradation thermique et la dégradation oxydative.

La dégradation thermique est comme son nom l'indique induite par la chaleur, mais est aussi impactée par la lumière et le pH (200,201). Les anthocyanes, sous leur forme chalcone, peuvent subir un clivage du cycle A et après réarrangement conduire à la formation d'un acide benzoïque et d'un benzaldéhyde (202). Cette réaction entraîne une décoloration irréversible des anthocyanes (203).

La dégradation oxydative implique quant à elle les anthocyanes et l'oxygène sous forme de peroxyde. Si les deux mécanismes de réaction sont différents, la dégradation oxydative conduit elle aussi à la formation d'un acide benzoïque et d'un benzaldéhyde, et à une décoloration irréversible (204–206).

Les produits de dégradation sont variés et dépendent de l'anthocyane en réaction. L'anthocyane principale des vins de *Vitis vinifera*, la malvidine-3-O-glucoside, se dégrade pour donner principalement l'acide syringique et le 2,4,6-trihydroxybenzaldéhyde (204,207).

## (2) *Copigmentation des anthocyanes*

En solution, les anthocyanes peuvent interagir avec d'autres molécules pour former des liaisons faibles de type  $\pi$ - $\pi$  *stacking* (208). Ces interactions impliquent deux cycles aromatiques qui, grâce à la délocalisation des électrons vont pouvoir s'associer en cumulant des interactions de Van der Waals et hydrophobes (209). Quand une anthocyane est impliquée, cela conduit à un effet hyperchrome (augmentation de l'intensité colorante) et bathochrome (augmentation de la longueur d'onde d'absorption)(210).

Dans les vins jeunes, cette interaction appelée copigmentation peut contribuer jusqu'à 50 % de la couleur (211,212). La plupart des composés phénoliques du vin sont répertoriés comme cofacteurs, bien que les principaux soient les anthocyanes elles-mêmes et les flavan-3-ols, en forte concentration dans le milieu (43,213,214). L'épicatchine particulièrement, du fait de sa conformation, permet de former des complexes plus efficacement que les autres composés (213,215).

Différents facteurs vont influencer la copigmentation et la force de la liaison, tels que le pH, le solvant, la concentration ou la température (216). Cependant, le facteur primordial reste la structure moléculaire du cofacteur qui va permettre l'accessibilité du cycle aromatique, sa réactivité, ou même la présence de groupements hydroxyles permettant des liaisons hydrogènes faibles renforçant l'interaction (217,218).

## (3) *Formation de pyranoanthocyanes*

La transformation des anthocyanes en composés plus stables pendant le vieillissement ne se limitent pas à leur condensation, elles peuvent aussi évoluer en pyranoanthocyanes. Ces composés se distinguent par la formation d'un nouvel hétérocycle de type pyrane, ce qui les rend insensibles à la décoloration au bisulfite de sodium (219,220). De plus, un effet hypsochrome est induit, qui diminue la longueur maximale d'absorption et se traduit par une évolution de la couleur du rouge vers un rouge-orangé (221).

Certaines des pyranoanthocyanes les plus importantes retrouvées dans les vins résultent de la réaction des anthocyanes monomères avec les métabolites de la levure libérées pendant la fermentation, telles que l'acide pyruvique, l'acide acétoacétique et l'acétaldéhyde (222). A noter que comme énoncé précédemment, l'acétaldéhyde peut aussi avoir comme origine l'oxydation de l'éthanol. La réaction de ces composés avec les anthocyanes va donner les pyranoanthocyanes de type vitisines, et est initiée par une substitution nucléophile sur le carbone 4 de l'anthocyane. Cette attaque va conduire à un

intermédiaire carbocation qui par réaction avec le groupement hydroxyle porté par le carbone 5, va former un nouvel hétérocycle (Figure 22).

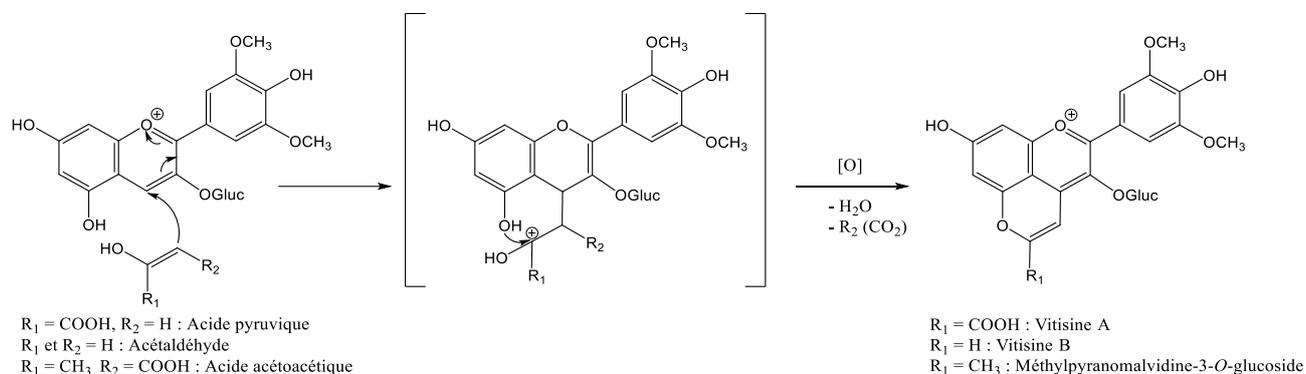


Figure 22 : Réaction de formations des pyranoanthocyanes de type Vitisines, exemple de la malvidine-3-*O*-glucoside, d'après de Freitas & Mateus, 2011

Les anthocyanes peuvent également réagir avec les acides hydroxycinnamiques pour former des hydroxypyrananthocyanes, avec les vinylflavan-3-ols issus du clivage d'un polymère lié par pont éthylidène pour former des flavano-pyranoanthocyanes, ou même avec une pyranoanthocyanine pour former des dimères (223–225)(Figure 23).

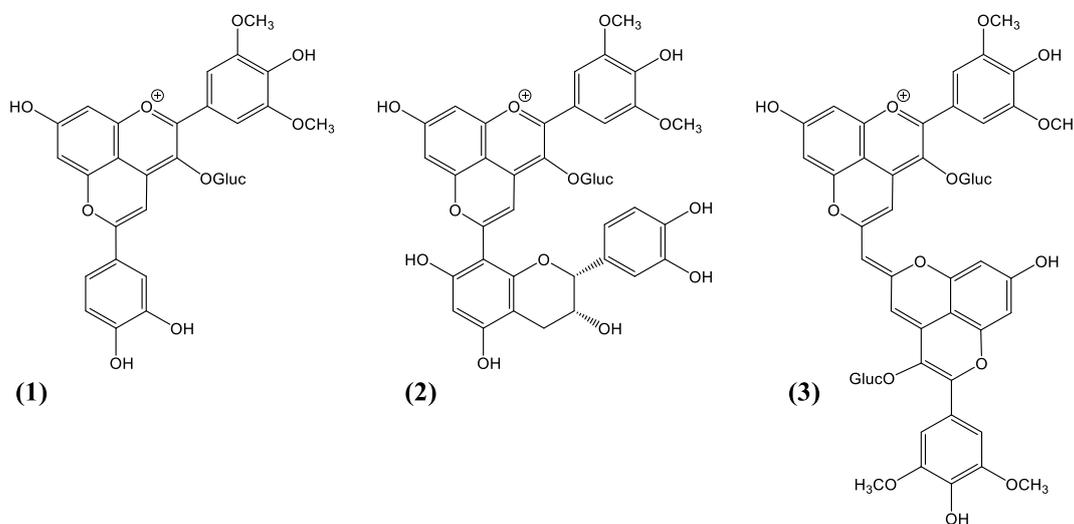


Figure 23 : Différentes structures des pyranoanthocyanes des vins : (1) exemple d'hydroxypyrananthocyanes, la pinotine A ; (2) flavano-pyranoanthocyanine issue de la réaction entre la malvidine-3-*O*-glucoside et l'épicatéchine ; (3) dimère de pyranoanthocyanine issu de la réaction entre deux malvidines-3-*O*-glucoside

Les pyranoanthocyanes retrouvées en plus grande concentration dans les vins sont les vitisines de type A, issues de la réaction avec l'acide pyruvique (225,226). Étant donné que les quantités les plus élevées d'acide pyruvique sont produites pendant la fermentation du vin rouge, la formation de ces

dérivés anthocyaniques est susceptible de commencer lors de ce processus (222). Au cours du vieillissement, les pyranoanthocyanes vont augmenter jusqu'à un maximum de quelques dizaines de mg/L après quatre ans, puis va décroître (199,225,227). Les pyranoanthocyanes possèdent une cinétique qui leur est propre et non corrélée à celle des anthocyanes libres. Cependant, comme les anthocyanes, leur concentration va être fortement impactée par le cépage utilisé (224).

#### ***d) Impact des macromolécules du vin sur les tanins condensés***

Les interactions intermoléculaires des proanthocyanidines sont basées sur plusieurs phénomènes, tous issus d'interactions électrostatiques qui conduisent à des liaisons faibles non covalentes. Ainsi, la structure chimique des tanins condensés les implique dans des liaisons hydrogènes grâce à la multitude de fonctions hydroxyles qu'ils présentent, et dans des liaisons de Van der Waals et hydrophobes grâce à leurs cycles aromatiques. Si ces forces sont plus faibles que les liaisons covalentes, elles ont un rayon d'action plus large, et permettent des interactions avec des composés macromoléculaires à fort encombrement stérique, comme les protéines ou les polysaccharides (48).

Les interactions des tanins condensés avec les protéines ont été largement étudiées, notamment à cause de leurs implications dans l'augmentation de la turbidité des vins, la précipitation du complexe tanin-protéine ou leur impact sur les perceptions d'astringence (228–230). Dans les vins, les protéines présentes naturellement viennent du raisin et des microorganismes (230,231). Elles peuvent aussi être ajoutées lors de l'étape de collage du vin afin de le clarifier, les plus fréquentes étant la gélatine, l'albumine et la caséine (232). Dans un vin fini, les concentrations en protéines oscillent entre 30 et 100 mg/L (231,233) Que leur origine soit endogène ou exogène, toutes les protéines retrouvées ont montré des capacités à précipiter préférentiellement les tanins condensés à fort poids moléculaires, et conduisent à une diminution de l'astringence (234,235). Cependant, cette baisse d'astringence ne semble pas être le résultat de la précipitation seule mais aussi celui de la formation de complexe tanin-protéine soluble dans le milieu (232).

Il a été démontré que les principaux polysaccharides du vin, notamment les mannoprotéines provenant de la dégradation des levures et des parois cellulaires végétales, interagissent aussi avec les tanins condensés (113). Elles se retrouvent en solution avec des concentrations oscillant entre 200 et 1500 mg/L (236). Plusieurs études ont montré l'importance des polysaccharides dans la médiation des interactions tanin-protéine, et leur impact sur la sensation d'astringence et la stabilisation de la couleur (237–239). Ainsi, certains polysaccharides exogènes, tel que la gomme arabique, peuvent

également être ajoutés au vin pour protéger la matière colloïdale et prévenir l'agrégation et la précipitation des complexes tanin-protéine (240,241).

De manière générale, les macromolécules du vin ont montré un impact non négligeable sur la concentration et la stabilité des proanthocyanidines en solution. De plus, l'utilisation de protéines pour le collage des vins, ainsi que de polysaccharides pour la stabilisation, implique des changements parfois importants dans la composition phénolique finale d'un vin, et doit être prise en compte.

## **5. Propriétés sensorielles des polyphénols du vin**

Les polyphénols, en particulier les anthocyanes et les tanins, sont recherchés dans les vins rouges pour les différentes propriétés sensorielles qu'ils amènent. Ils sont ainsi majoritairement responsables de la couleur, de l'astringence et de l'amertume des vins.

Le premier contact du consommateur avec le vin se fait au travers de sa vue, et sa couleur influence la perception aromatique et gustative (242,243). La couleur a donc un impact très important sur la consommation, et il a été démontré qu'il est possible de leurrer des dégustateurs entraînés en changeant artificiellement la couleur d'un vin (244). Contrairement au goût ou à l'odorat, parfois difficiles à appréhender chimiquement, la couleur peut aisément s'expliquer par la capacité d'une molécule à absorber ou non les différentes longueurs d'onde du visible. Ainsi, comme il a été détaillé précédemment, ce sont les anthocyanes qui sont responsables de la couleur des vins rouges, et leur dégradation, polymérisation ou transformation sont responsables de l'évolution de la couleur lors du vieillissement.

L'astringence et l'amertume sont quant à elles des propriétés organoleptiques plus complexes à appréhender, et les interactions ou réactions chimiques qui les induisent restent souvent mal comprises. Cependant le rôle prédominant des tanins condensés dans ces perceptions a pu être démontré à plusieurs reprises (245,246).

### ***a) Perception d'astringence***

L'astringence peut être décrite par les sensations complexes dues aux rétrécissements, aux étirements ou aux plissements de l'épithélium de la bouche suite à l'exposition de substances telles que les sels d'aluminium ou les tanins (247). De manière plus spécifique, elle est décrite par les consommateurs de vin comme la sensation de sécheresse, de rugosité et de manque de lubrification en bouche (248). Cette sensation n'est pas spécifique à une région de la bouche, et est plus considéré

comme un stimulus tactile (249,250). De plus, lors d'ingestions répétées, la perception de l'astringence subit un phénomène d'addition et une augmentation de l'intensité perçue (251).

Le mécanisme d'action entraînant cette sensation fait consensus, et vient de la capacité des tanins à précipiter les protéines salivaires. Les groupements hydroxyles des composés phénoliques viennent former des liaisons hydrogènes avec les groupements carboxyles et amides des protéines, et des interactions hydrophobes se mettent en place entre les cycles aromatiques des polyphénols et les chaînes latérales apolaires d'acides aminés tels que la leucine, la lysine ou la proline (252–254)(Figure 24).

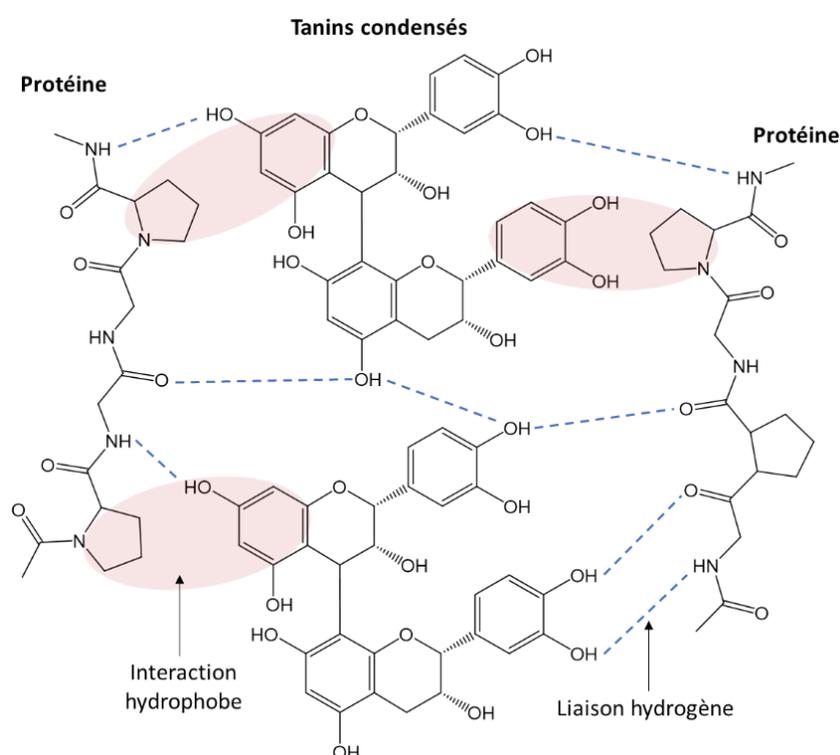


Figure 24 : Schéma des interactions entre tanins condensés et protéines

Plus spécifiquement, les protéines riches en proline, comme les protéines de la salive, sont de très bons sites d'interactions. Elles fournissent grâce à leur hétérocycle aminé une surface plane, rigide et hydrophobe extrêmement favorable aux interactions avec les cycles aromatiques des tanins condensés (255).

### ***b) Perception d'amertume***

L'amertume, contrairement à l'astringence, fait partie des cinq saveurs primaires avec le sucré, le salé, l'acide et l'umami. Elle se démarque cependant des autres saveurs par un temps plus

long pour atteindre son intensité maximale en bouche (256). Parfois difficile à apprécier, l'amertume est considérée comme un mécanisme de défense contre les substances toxiques, et active ainsi des signaux d'avertissement pour réguler leur ingestion (257). La construction de la perception du goût amer est un processus qui varie en fonction des individus, et dépend des récepteurs chimiques individuels de l'amertume (258).

Le ressenti de l'amertume se fait au travers de 25 récepteurs transmembranaires spécifiques (TAS<sub>2</sub>R), préférentiellement situés sur la partie centrale arrière de la langue (259,260). Lors de la consommation de vin, les composés amers induisent un changement conformationnel de ces récepteurs, qui après plusieurs réactions intracellulaires complexes va conduire à la libération des neurotransmetteurs vecteurs de la perception (261,262). Les agonistes amers ont une grande diversité structurale, comprenant des peptides, des phénols, des acides aminés ou encore des lactones, et les récepteurs TAS<sub>2</sub>R sont généralement spécifiques à un groupe de composés, ce qui les rendent parfois sensibles à plusieurs molécules (256). Cependant, le mécanisme de ces spécificités est encore mal connu.

### ***c) Astringence et amertume des polyphénols du vin***

Si dans les vins rouges, les perceptions d'astringence et d'amertume sont induites majoritairement par la teneur en tanins condensés, d'autres facteurs vont les influencer (245,263,264). Le degré de polymérisation, la composition en unités flavan-3-ols des chaînes proanthocyanidiques ainsi que les variabilités de la matrice du vin vont ainsi jouer un rôle, et faire fluctuer les propriétés organoleptiques des tanins condensés (246).

Le degré de polymérisation moyen joue un rôle prépondérant dans la perception d'astringence et d'amertume. Il a ainsi été démontré que les monomères proposent une intensité de perception plus faible que les polymères. L'astringence et l'amertume vont augmenter avec le degré de polymérisation jusqu'à un maximum, puis diminuer (3,246). D'après les travaux de Lea et al., la perception d'astringence atteint son maximum pour les heptamères de tanins condensés, alors que pour l'amertume il s'agit du tétramère. Les monomères et les tanins faiblement condensés sont perçus plus amers qu'astringents, et cette tendance s'inverse quand le degré de polymérisation dépasse 6 (265)(Figure 25).

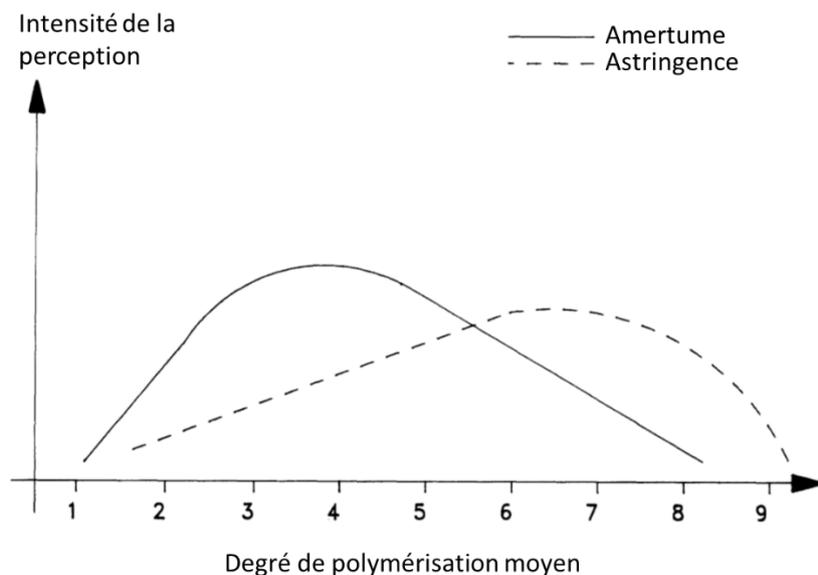


Figure 25 : Intensité de la perception d'astringence et d'amertume en fonction du degré de polymérisation moyen des tanins condensés (265)

Ces observations ont été corroborées par l'analyse de l'affinité chimique des tanins avec les protéines salivaires, et leur faculté à se lier en fonction du degré de polymérisation (266). En présence d'un excès de protéines salivaires, les monomères ne présentent pas de réactivité. Celle-ci apparaît à partir des dimères, et augmente de manière linéaire jusqu'à une précipitation totale pour les pentamères. Si l'évolution de la perception des structures de proanthocyanidines jusqu'à un degré de polymérisation de 10 semble faire consensus, l'impact des grandes structures reste à éclaircir. Dans le passé Lea, 1992, pensait que ces structures n'étaient pas assez solubles pour être astringentes. Cependant, plus récemment, Cheynier et al., 2006, ont travaillé avec des structures hautement polymérisées de proanthocyanidines (Degré de polymérisation moyen jusqu'à 70) et ont observé qu'elles étaient solubles en solution hydroalcoolique et hautement astringentes (1).

La structure des flavan-3-ols va aussi impacter les perceptions. Il a notamment été montré que l'épicatéchine était perçue plus astringente et amère que la catéchine (267,268). De plus, l'augmentation du pourcentage de galloylation induit une augmentation de la perception, ce qui rend les tanins issus des pépins plus astringents et amers que ceux de la pellicule, indépendamment de leur degré de polymérisation (268–271).

Des paramètres extérieurs à la composition phénolique peuvent aussi agir sur les perceptions des tanins du vin. L'augmentation des teneurs en acide tartrique, par exemple, a montré une corrélation positive avec l'augmentation de l'astringence (264). A contrario, une augmentation du pH ou de la

proportion d'éthanol va diminuer sa sensation (264,272). Concernant l'amertume des tanins, ces variables ne semblent pas la moduler, à l'exception de l'éthanol qui induit une augmentation de sa perception (272). La concentration en polysaccharides et en protéines est aussi un facteur à intégrer. Comme évoqué précédemment, ces composés peuvent agir sur la précipitation des tanins et ainsi diminuer l'astringence et l'amertume (232,238).

Très récemment, il a aussi été démontré que les anthocyanes libres réagissaient avec les protéines de la salive et ont pu être décrites en dégustation comme astringentes et amères (273). Leur concentration est donc également à prendre en compte dans l'appréciation globale de l'astringence et l'amertume des vins.

## **B. La spectroscopie UV-visible et IRTF appliquée aux polyphénols du raisin, du moût et du vin**

Un nombre important de méthodes ont été développées pour doser et identifier les différents polyphénols du vin. Elles peuvent se fonder sur les propriétés physicochimiques de ces composés, comme la transformation de proanthocyanidines et anthocyanidines (114), l'action décolorante du bisulfite de sodium sur les anthocyanes (274), ou encore leur faculté à précipiter avec les protéines (275). L'avantage de ces techniques c'est qu'elles mesurent une classe entière de composés, et non un seul marqueur moléculaire, et permettent ainsi d'obtenir une vision globale. Plusieurs méthodes de routine ont été développées au fil du temps pour des analyses plus fines des polyphénols en utilisant la chromatographie liquide haute performance couplée à la spectrométrie de masse, et permettent de doser les polyphénols principaux du vin, tels que les anthocyanes, les flavan-3-ols, les ellagitannins, etc (163,276,277). Cependant, toutes ces techniques demandent du temps d'analyse, des compétences et connaissances spécifiques pour être appliquées, ainsi que du matériel souvent extrêmement coûteux. Pour toutes ces raisons, elles ne sont pas adaptées à une analyse de routine rapide en milieu industriel.

Pour faciliter ces analyses il est possible d'utiliser la spectroscopie. Le principe de cette technologie s'appuie sur la capacité d'une molécule à absorber une quantité spécifique d'un rayonnement électromagnétique d'une longueur d'onde donnée. Les longueurs d'ondes d'absorbances permettront ainsi de caractériser le composé. La valeur d'absorbance, quant à elle, suit la loi de Beer Lambert, c'est-à-dire qu'elle est corrélée de façon linéaire avec la concentration, et permet donc de quantifier

le composé (278). Pour l'analyse des polyphénols, deux zones spectrales se détachent alors, l'UV-visible et l'infrarouge (12,279).

### **1. La spectroscopie UV-visible appliquée aux polyphénols du raisin, du moût et du vin**

La spectroscopie UV-visible se décompose en deux zones spectrales. Tout d'abord l'ultraviolet, appelé aussi lumière noire (car invisible à l'œil nu), est un rayonnement électromagnétique qui couvre les longueurs d'ondes comprises entre 10 et 400 nm. Cependant, en dessous de 200 nm, les rayons ultraviolets ne se propagent plus dans l'air, la spectroscopie UV classique s'arrête donc à cette longueur d'onde. L'absorption dans l'UV est très spécifique aux composés insaturés, et permet ainsi de détecter les doubles liaisons et les liaisons  $\pi$  (11). La seconde zone spectrale est le visible, il s'agit d'un rayonnement électromagnétique qui est constitué du spectre de la lumière visible, compris entre 400 et 800 nm. L'absorption dans cette zone induit alors le caractère coloré de la molécule étudiée (280).

La technologie UV-visible semble donc très adaptée à la quantification des polyphénols. Elle est d'ailleurs déjà fortement utilisée pour déterminer l'indice de phénols totaux ainsi que l'intensité colorante des vins (281). Les cycles phénoliques de ces composés, fortement insaturés, vont absorber la lumière UV, alors que les anthocyanes et leurs dérivés vont absorber dans le visible (280,282). De plus, les composés prédominants de la matrice du vin et du moût, tels que les sucres, l'alcool ou encore les acides non phénoliques ne présentent pas de signaux significatifs en UV-visible (11). La réponse spectrale entre 200 et 800 nm est donc majoritairement due à la composition phénolique. Il faut noter cependant que suivant les molécules en présence dans le milieu, le spectre UV-visible général va différer car chaque composé phénolique a ses propriétés d'absorption et donc son empreinte spectrale. La Figure 26 montre ainsi que certaines structures proches comme la malvidine-3-*O*-glucoside et la malvidine-3-*O-p*-coumaroylglucoside présentent des bandes d'absorption principales presque identiques, mais que les intensités peuvent varier pour conduire à une différenciation spectrale marquée. A contrario, des composés comme la catéchine et l'acide gallique, dont la structure moléculaire est bien plus différenciée, présentent des spectres relativement simples avec un maximum d'absorption très proche qui rend la distinction plus difficile.

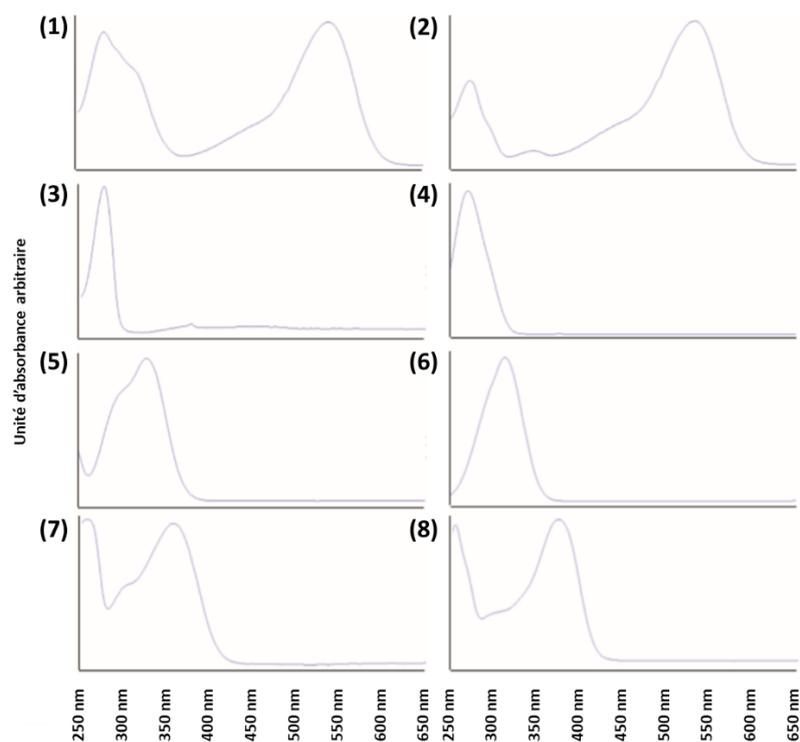


Figure 26 : Spectres de différents composés phénoliques : (1) malvidine-3-*O*-glucoside ; (2) malvidine-3-*O*-*p*-coumaroylglucoside ; (3) catéchine ; (4) acide gallique ; (5) acide caftarique ; (6) acide coutarique ; (7) quercétine-3-*O*-diglucoside ; (8) quercétine (283)

### a) La spectroscopie UV appliquée aux tanins condensés

Les flavan-3-ols présentent tous une bande d'absorption principale autour de 280 nm, comme présenté sur le spectre de la catéchine de la Figure 26 (282). Les tanins condensés sont les principaux composés phénoliques du vin, ce qui explique notamment pourquoi cette longueur d'onde est utilisée comme un indice de phénols totaux (281). Si la longueur d'onde du maximum d'absorbance reste inchangée, certaines différences sont notables sur le spectre global des différents monomères de flavan-3-ols. Ainsi, la catéchine et l'épicatéchine ont des spectres relativement similaires, alors que l'épicatéchine-gallate montre un épaulement du spectre autour de 310 nm (283). De plus, à concentration identique, l'épicatéchine-gallate montre une absorbance deux fois supérieure à celle de la catéchine et l'épicatéchine (282)(Tableau 4). A contrario, l'absorbance de l'épigallocatechine apparaît très faible en comparaison. Le degré de polymérisation moyen, quant à lui, ne semble pas impacter de manière significative l'absorbance globale.

Composé à 1 mg/L	Absorbance à 280 nm
Catéchine	0,012
Epicatéchine	0,014
Epigallocatechine	0,004
Epicatéchine-gallate	0,032
Malvidine-3- <i>O</i> -glucoside	0,026
Acide gallique	0,049
Acide caftarique	0,02
Quercétine	0,0016
Tanins de pomme (DPm = 7)	0,013
Tanins de pomme (DPm = 12)	0,014

Tableau 4 : Absorbance à 280 nm de plusieurs polyphénols du vin, dilué à 1 mg/L dans une solution à 10 % de méthanol et 1 % d'acide chlorhydrique (282) (DPm : degré de polymérisation moyen)

Si l'UV démontre de réelles qualités pour quantifier les polyphénols, il montre aussi ses limites lorsque l'on se restreint à une lecture directe de l'absorbance pour une longueur d'onde donnée. Les spectres UV de nombreux composés vont se chevaucher, s'additionner, et la variabilité de l'absorbance est issue potentiellement d'origines multiples. En prenant l'exemple de l'absorbance à 280 nm, bien que la majorité de la réponse soit induite par les tanins, les anthocyanes ou les acides du vin peuvent aussi impacter la lecture (282).

### ***b) La spectroscopie visible appliquée aux anthocyanes***

Dans le visible, le nombre de familles de composés possédant une absorbance est bien plus réduit, et se limite principalement aux anthocyanes et à leurs dérivés. Tout comme pour les flavan-3-ols, la nature du monomère va influencer la réponse spectrale, impactant la longueur d'onde maximale d'absorbance, ainsi que le coefficient d'extinction (Tableau 5). Le maximum d'absorbance des monomères d'anthocyanes se situe aux alentours de 520 nm, ce qui correspond à un rayonnement de couleur verte. Le fait que les anthocyanes absorbent la composante verte du spectre lumineux explique donc leur couleur rouge caractéristique perçue par l'œil humain. La variabilité du coefficient d'extinction apparaît moins importante que pour les flavan-3-ols, avec deux monomères qui se démarquent. A concentration égale, la delphinidine-3-*O*-glucoside sera donc l'anthocyane qui présentera la plus grande absorbance, alors que la péonidine-3-*O*-glucoside sera celle qui absorbera le moins. Cependant, contrairement aux flavan-3-ols, des paramètres de la matrice du vin comme le pH joue un rôle important sur l'équilibre des anthocyanes, et vont donc aussi impacter l'absorbance (128).

<b>Anthocyane</b>	<b><math>\lambda_{\text{visible max}}</math></b>	<b>Coefficient d'extinction <math>\text{L.mol}^{-1}.\text{cm}^{-1}</math></b>
Cyanidine-3- <i>O</i> -glucoside	512	18800
Delphinidine-3- <i>O</i> -glucoside	520	23700
Péonidine-3- <i>O</i> -glucoside	512	14100
Pétunidine-3- <i>O</i> -glucoside	520	18900
Malvidine-3- <i>O</i> -glucoside	520	20200

Tableau 5 : Longueurs d'ondes maximales d'absorbance et coefficients d'extinction molaire des cinq principales anthocyanes monoglucosidées du vin, en solution aqueuse à 10 % d'éthanol et pH 1,5 (128)

L'évolution des anthocyanes a aussi un fort impact sur la couleur des vins, et donc sur l'absorbance dans le visible. Comme cela a pu être évoqué précédemment, la complexation des anthocyanes avec les ellagitanins ainsi que la copigmentation va induire un effet bathochrome, et donc déplacer le maximum d'absorption vers des longueurs d'ondes plus élevées avec pour effet d'accentuer la composante bleue-mauve (176,210). A l'opposé, la condensation d'une anthocyane avec des tanins ou sa transformation en pyranoanthocyanes va induire un effet hypsochrome, et donc déplacer le maximum d'absorption vers des longueurs d'ondes plus faibles avec pour effet d'accentuer la composante orangée (221). La spectroscopie visible semble donc tout indiquée pour suivre l'évolution des anthocyanes du vin.

## **2. L'infrarouge à transformée de Fourier appliqué aux polyphénols du raisin, du moût et du vin**

La spectroscopie infrarouge utilise les propriétés intrinsèques des molécules à absorber des rayonnements électromagnétiques à des fréquences spécifiques, en fonction de la nature des liaisons moléculaires qui les composent. Le rayonnement infrarouge absorbé va avoir pour effet d'exciter la liaison entre deux atomes et va engendrer la vibration de celle-ci. Deux types vibrations sont alors possibles : une vibration d'étirement, où la vibration s'effectue dans l'axe de la liaison et influe sur sa longueur, et une vibration de déformation, où la vibration s'effectue en dehors de l'axe de la liaison sans influencer sa longueur (Figure 27).

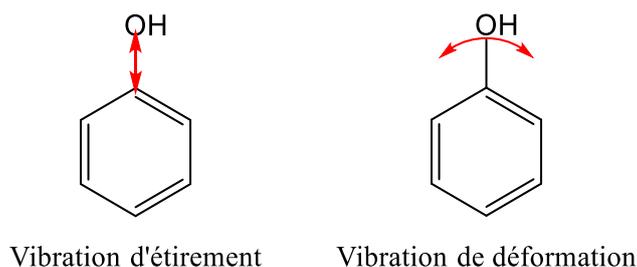


Figure 27 : Différents types de vibrations induits par un rayonnement infrarouge

Le terme infrarouge se réfère généralement à tout rayonnement électromagnétique entre 0,7 mm et 1000 mm. Néanmoins, la région de l'infrarouge moyen, comprise entre 2,5 mm et 25 mm (soit entre 4000 et 400  $\text{cm}^{-1}$ ), est la plus attractive pour la chimie analytique. Cette zone spectrale comprend les fréquences fondamentales de la majorité des composés organiques, avec des pics d'absorbance généralement étroits et distincts. Ces caractéristiques permettent de faire des étalonnages qualitatifs simples et robustes, et largement insensibles aux artefacts (284).

Il est possible d'enregistrer un spectre infrarouge longueur d'onde par longueur d'onde, comme cela est fait notamment pour la spectroscopie UV-visible. Cependant, l'infrarouge dispose d'une large bande spectrale qui demanderait un temps d'analyse important. Des outils mathématiques ont donc été développés pour réduire drastiquement le temps d'acquisition des spectres, et ont conduit à l'infrarouge à transformée de Fourier (IRTF). La puissance de l'analyse de Fourier est de pouvoir décomposer un signal en composantes de fréquence individuelles et d'établir l'intensité relative de chaque composante (285). Cette transformation permet de représenter une fréquence (définie par un nombre d'occurrences dans le temps) par un spectre.

En pratique, au lieu d'utiliser une unique longueur d'onde, l'échantillon est exposé à tout le spectre. Ce rayonnement passe par un interféromètre qui permet de réduire le spectre à plusieurs longueurs d'ondes définies, et cette opération est répétée pour couvrir toutes les longueurs d'ondes. L'analyse par transformée de Fourier permet alors de traiter l'interférogramme obtenu afin d'obtenir le spectre correspondant (286).

### ***a) L'infrarouge à transformée de Fourier appliquée aux polyphénols***

L'utilisation de la spectroscopie infrarouge à transformée de Fourier est récemment apparue comme une solution possible pour la mesure rapide des tanins du vin (287–289). C'est dans l'infrarouge moyen que la réponse spectrale sera la plus importante, ce qui en fait une zone d'intérêt pour la mesure des tanins. Du fait de la complexité de ces molécules, les spectres obtenus sont bien

plus complexes qu'en spectroscopie UV-visible, car chaque type de liaison possède son absorbance propre. Les vibrations de déformation et d'étirement n'absorbant pas à la même longueur d'onde, les tanins condensés présentent une multitude de pics caractéristiques. La Figure 28 montre le spectre d'absorbance de tanins commerciaux de raisins à différentes concentrations (15).

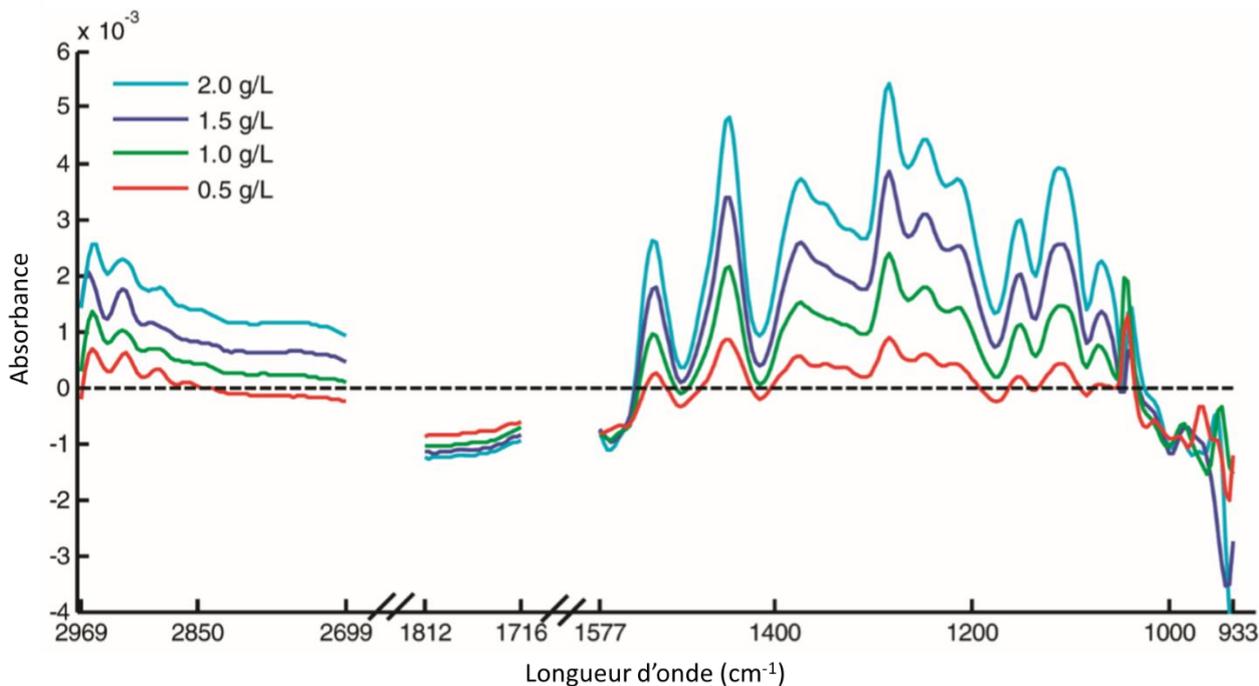


Figure 28 : Spectre infrarouge moyen des tanins du vin (15)

Les spectres présentés indiquent une absorbance corrélée à la concentration, idéale pour la quantification. Cette réponse est notamment très lisible dans la région de 2969 à 2699  $\text{cm}^{-1}$  ainsi que de 1577 à 1060  $\text{cm}^{-1}$ . Aujourd'hui, la plupart des longueurs d'ondes spécifiques aux polyphénols du vin ont été identifiées, et sont indiquées dans le Tableau 6.

Vibration des groupes fonctionnels	Région de la réponse spectrale (cm <sup>-1</sup> )
<b>Cycle phénolique</b>	
Étirement C=C	1640-1430
Déformation O-H	1410-1260
Étirement C-O	1150-1040
Déformation C-O-H	1390-1310
Déformation dans le plan =CH	1290-1000
Déformation hors du plan =CH	900-700
Étirement liaison C-C	1225-1075
Harmonique aromatique et combinaison	2000-1700
<b>Acide phénol</b>	
Étirement C=O	1755-1630
Déformation O-H	1380-1280
Étirement C-O	1190-1075
Déformation hors du plan =CH	990-960
Déformation C-H (méthoxy)	1470-1435
Déformation C-O (méthoxy)	1310-1210, 1200-1185, 1120-1020
<b>Flavonoïde</b>	
Étirement C-H (alcane)	3000-2840
Étirement C=O	1670-1625
Combinaison étirement C=O et C=C	1570-1525, 1465-1445, 1420-1400
Déformation dans le plan du cycle pyrylium	1650-1615, 1560-1465, 1450-1400, 1000-970
Déformation hors du plan du cycle pyrylium	920-870
Étirement C-O (éther)	1310-1230
Étirement C-O (glucoside)	1200-1030
Étirement C-O-C	1110-1090, 820-805
<b>Proanthocyanidine</b>	
Étirement C-O	1300-1100
Étirement symétrique C-O-C	1081
Déformation O-H	868
Torsion aromatique et déformation C-H	950-700
<b>Stilbène</b>	
Étirement C=C ( <i>trans</i> -)	1670-1610
Déformation dans le plan C-H	1330-1215
Déformation C=C ( <i>trans</i> -)	980-955

Tableau 6 : Principales réponses spectrales des groupes fonctionnels composant les polyphénols du vin (290–292)

La réponse spectrale dans la zone de 3000 à 2800 cm<sup>-1</sup> est due à l'étirement des liaisons entre carbone et hydrogène des alcanes. Ainsi, pour la mesure des polyphénols, elle peut être exclusivement

attribuée aux flavonoïdes, et plus généralement aux flavan-3-ols qui sont la classe de molécules très majoritaires de cette famille. Il s'agit donc d'une zone d'intérêt pour la quantification des tanins du vin (15). Cependant, d'autres molécules du vin possèdent aussi des absorbances à ces longueurs d'ondes, comme l'éthanol ou l'acide tartrique, et cette zone seule n'est pas suffisante pour caractériser les tanins. Dès lors, l'utilisation du spectre complet en infrarouge moyen semble primordiale. La zone du spectre entre 1600 et 800  $\text{cm}^{-1}$  se montre extrêmement riche en informations, mais un grand nombre de vibrations différentes peuvent se superposer et s'additionner et interférer sur la lecture (290). Il est cependant possible de différencier l'empreinte spectrale des différents composés, comme cela a pu être démontré par Foo en 1981. Il a ainsi prouvé que la réponse spectrale en infrarouge moyen est différente pour chacun des monomères de flavan-3-ols, et que cette différence est maintenue lors de la condensation de ces derniers (293). Le spectre d'un dimère de catéchine va différer légèrement d'un dimère d'épicatéchine, alors que la présence de gallocatéchine va avoir une influence bien plus importante sur le spectre global.

L'IRTF, tout comme l'UV-visible, apparaît comme une technique d'analyse rapide et efficace pour la caractérisation des polyphénols du vin, et plus précisément des tanins et des anthocyanes. Si la théorie montre l'intérêt de ces technologies sur des molécules ou produits purs, en pratique, la complexité de la matrice du vin ne permet pas une lecture directe efficace. De nombreux composés vont venir perturber la réponse spectrale et empêcher la possibilité d'une caractérisation et quantification. Deux solutions apparaissent donc pour éviter ces interférences. La première est de purifier l'échantillon afin de ne garder que les polyphénols, mais va à l'encontre d'une analyse rapide et facile à mettre en œuvre. La seconde solution est d'utiliser l'analyse de données multivariée pour réussir à développer des modèles de prédiction des composés d'intérêts, et s'affranchir de l'impact de la matrice (6,279,287).

### **C. L'analyse de données multivariée : généralités et application aux polyphénols du raisin, du moût et du vin**

L'analyse statistique multivariée, aussi appelée chimiométrie, permet d'extraire les informations contenues dans un spectre complexe (6). Elle se base essentiellement sur des techniques de régression qui sont généralement couplées à un prétraitement de l'information spectrale (294–296).

Des applications tant quantitatives que qualitatives de l'analyse de données multivariée sont rapportées dans la littérature pour l'analyse du raisin et du vin (297–300). Il s'agit d'une méthode de modélisation dite « douce » où l'accent est mis sur les propriétés du signal plutôt que sur le bruit et la structure, ce qui l'oppose aux statistiques (8,301,302). La supposition de base est que le système chimique étudié est complexe et qu'il n'est alors pas possible de décrire correctement son comportement à l'aide d'un modèle statistique classique. La méthode s'appuie donc sur la variation et la corrélation des données analysées, qui permettent l'obtention d'une matrice de covariance. Cette matrice peut être considérée comme une mesure de la fluctuation et des interactions de chaque variable indépendante présente dans l'ensemble des données (303,304).

Plusieurs algorithmes ont été développés pour permettre d'utiliser l'analyse de données multivariée pour la quantification, la caractérisation et la discrimination de composés à l'aide de méthode spectroscopique. Le plus performant et le plus couramment utilisé aujourd'hui est la régression PLS (Partial Least Squares regression) appelé aussi régression par les moindres carrés partiels (8,296,305).

## **1. La régression PLS (régression par les moindres carrés partiels)**

### ***a) Description de l'analyse et la régression PLS***

L'analyse PLS est une généralisation de l'analyse linéaire multiple qui, contrairement à cette dernière, permet d'analyser des données avec des variables fortement corrélées, bruyantes et redondantes (8,296,306). Les longueurs d'ondes d'un spectre chimique sont le parfait exemple de ce genre de variables. Cette approche a été développée vers 1975 par Herman Wold pour pouvoir modéliser des données complexes (8). Le modèle global de l'analyse PLS est décrit par deux modèles sous-jacents. Le premier est un modèle de mesure qui permet de relier les variables observées à des variables dites latentes, qui ne sont pas observables. Le second est un modèle structurel qui va relier ces variables latentes à une ou plusieurs nouvelles variables latentes, en tentant d'établir des liens décrits par des équations linéaires.

Les travaux de Tenenhaus et al., (2005) permettent d'expliquer ce fonctionnement hiérarchique de l'analyse PLS. On considère qu'un jeu de données est expliqué par  $J$  blocs de variables  $X_1, X_2, \dots, X_J$ . Chaque bloc de variables peut être expliqué par une variable latente dite endogène  $\xi_1, \xi_2, \dots, \xi_J$ , qui résumera la variabilité du bloc (Figure 29). Les relations structurelles de ces variables endogènes permettront alors de modéliser une nouvelle variable latente  $\xi$  dite exogène qui cherchera à représenter toute la variabilité du jeu de données.

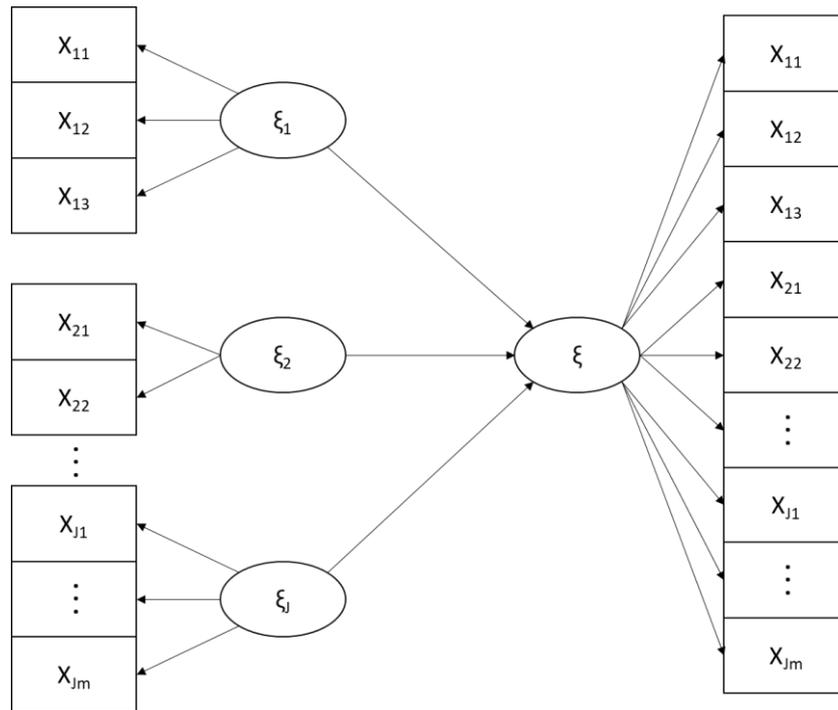


Figure 29 : Modèle hiérarchique pour une analyse PLS de J blocs de variables (302)

Le cas présenté ici montre la variabilité d'un jeu de données expliquée par une seule variable latente, aussi appelée composante, mais en pratique les données étudiées sont plus complexes et en nécessitent plusieurs (302,307,308). Il faut cependant être vigilant dans l'utilisation d'un nombre élevé de variables latentes lorsque le nombre d'individus étudiés est faible, car cela peut conduire à une surestimation de la modélisation de la variabilité (279). A contrario, l'utilisation d'un nombre trop faible de composantes entraînera une sous-estimation du modèle qui ne sera pas assez robuste pour expliquer au mieux la variabilité des données.

La régression PLS permet d'utiliser l'analyse PLS sur deux matrices de variables distinctes X et Y afin de rechercher les corrélations possibles entre elles. Si l'on applique cette méthode à la chimie, la matrice X représentera les propriétés chimiques de l'ensemble des échantillons, et Y leur composition chimique (8). Ainsi, la matrice X peut être produite par l'analyse spectrale IRTF ou UV-visible, et la matrice Y par l'analyse de référence du composé recherché. La régression PLS va donc créer un modèle structurel qui recherchera la corrélation entre les deux matrices, et permettra, si l'information est exploitable, de prédire l'analyse de référence grâce à l'analyse spectrale.

Il est aussi possible d'utiliser la régression PLS pour de l'analyse discriminante, ce qui permet soit de définir des catégories d'individus, ou de définir une démarcation entre individus. C'est utile par exemple pour déterminer si un échantillon est considéré comme bon ou mauvais, ou s'il appartient à

une classe particulière (6). Un modèle de classification peut ainsi être construit pour prédire une classe d'échantillons en comparant l'échantillon à un ensemble d'expériences précédemment analysé, dans lequel les catégories d'échantillons sont déjà connues.

### ***b) Prétraitement des données spectrales***

Il arrive que les informations pertinentes d'un spectre ne puissent pas être modélisées de manière efficace par les techniques de régression linéaire, il est alors nécessaire de prétraiter les données. Plusieurs méthodes ont donc été développées dans ce but (296,309). Lorsque les spectres mesurés sont sujets aux bruits de fond, à une ligne de base ou d'autres facteurs indésirables, le prétraitement devient dès lors indispensable (310).

Le prétraitement le plus utilisé est le centrage par la moyenne (*mean centering*), et consiste à soustraire la moyenne d'une variable de toutes les observations sur cette variable dans l'ensemble de données. Ainsi, la nouvelle moyenne de la variable est égale à zéro (311). Ce traitement est particulièrement utile pour éviter les interférences de colinéarités multiples non essentielles, qui peuvent notamment être induites par le bruit de fond. Le centrage par la moyenne s'assure que toutes les variables soient interprétables en termes de variation, et est fortement recommandé pour toutes applications pratiques (312). Certaines transformations sont aussi nécessaires, notamment pour l'étude de spectres IRTF. Les instruments de spectroscopie enregistrent les données en transmittance, il est alors indispensable de les transformer en absorbance pour que la loi de Beer Lambert s'applique (312).

Plusieurs autres techniques de prétraitement plus complexes ont aussi été développées. Le lissage par exemple (*smoothing*), est proposé pour supprimer le bruit aléatoire. Si cela améliore l'aspect visuel des spectres, le lissage supprime aussi des informations et n'est pas forcément recommandé (296). Il en va de même pour des techniques de standardisation, normalisation ou plus complexes encore, de régression linéaire et non linéaire (312). Si le prétraitement des données spectrales est un outil puissant pour améliorer la prédiction des modèles développés, chaque transformation doit être appliquée en connaissance de cause pour éviter une surinterprétation des résultats ou à l'inverse, de perdre l'information (6).

### ***c) Etapes de construction d'un modèle de prédiction***

La construction d'un modèle de prédiction pour un composé d'intérêt ou pour la discrimination d'échantillons se décompose en quatre étapes : la sélection de la base d'échantillons

qui va servir à la calibration, l'analyse de référence qui va être modélisée, l'analyse spectrale et pour finir l'analyse statistique multivariée pour concevoir le modèle. Dès lors, plusieurs paramètres sont à prendre en compte pour obtenir le modèle le plus performant possible et éviter certains biais de prédiction.

La sélection des échantillons pour l'étalonnage est primordiale. Ils doivent être sélectionnés parmi des échantillons similaires à ceux qui vont être analysés dans le futur, et subir les mêmes traitements (313,314). De plus, il est nécessaire qu'ils couvrent une large gamme de concentrations du composé étudié, en essayant au maximum d'avoir une répartition homogène des teneurs. Pour finir, il est fortement conseillé d'avoir une forte variabilité de la matrice, afin que le modèle construit soit adapté à tout type d'échantillons (315).

Il faut comprendre et maîtriser l'analyse de référence utilisée, afin d'en connaître les avantages et les limites. Des paramètres tels que l'erreur standard de la méthode vont influencer la robustesse du modèle, celui-ci ne pouvant au maximum qu'égaliser la précision de l'analyse de laboratoire (6,316).

L'interprétation des résultats de calibration couplée à une connaissance de la technologie spectrale utilisée est aussi à prendre en compte. Comme l'analyse multivariée va rechercher la meilleure corrélation possible entre l'analyse de référence et les données spectrales sans discernement, c'est à l'utilisateur de s'assurer que les résultats obtenus ne sont pas le fruit d'une corrélation induite par un autre composé (6). L'exemple le plus commun serait d'attribuer la diminution du taux de sucre des moûts pendant la fermentation alcoolique à l'augmentation des teneurs en éthanol, et donc d'obtenir une corrélation indirecte qui n'est pas construite sur la réponse spectrale du composé étudié.

#### ***d) Validation d'un modèle de prédiction***

Afin de vérifier la précision d'un modèle de prédiction et éviter une surestimation de sa robustesse, des procédures de validation doivent être appliquées. En effet, il sera difficile de juger un modèle d'étalonnage sans une validation appropriée. Si celui-ci peut être capable de représenter correctement la variabilité du jeu de données, cela n'implique pas forcément une capacité à prédire des échantillons extérieurs (279).

La méthode la plus efficace pour valider un modèle est d'utiliser un jeu de données indépendant qui servira de set de validation (317). Cependant, la conception d'un modèle de prédiction par régression PLS demande l'utilisation d'un nombre important d'échantillons pour que la calibration soit performante, et le set de validation doit lui aussi contenir un nombre suffisant d'échantillons pour être

pertinent. En pratique, il est souvent préféré à cette méthode une option plus facile à mettre en œuvre, la validation croisée.

Il existe plusieurs méthodes de validation croisée, celle étant le plus régulièrement utilisée est la *leave-one-out cross-validation* ou validation croisée d'un contre tous. Un échantillon est retiré du jeu de données et le modèle de calibration est construit à partir des données restantes (296,309). L'échantillon retiré est alors utilisé comme validation. Ce processus est réitéré afin que tous les échantillons aient servi de validation (Figure 30).

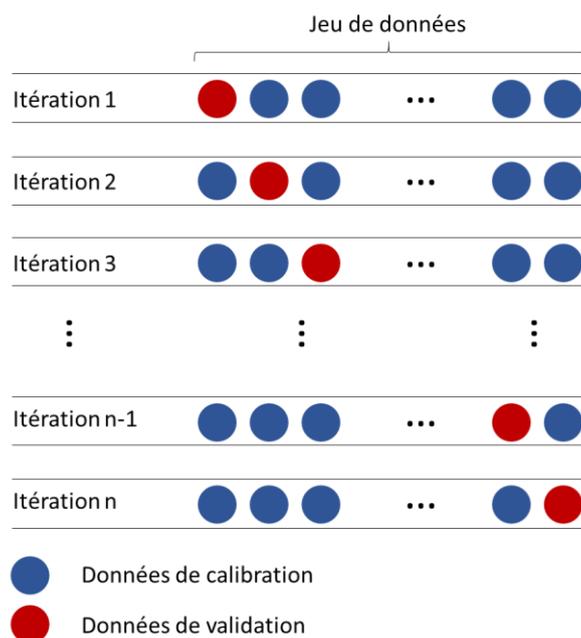


Figure 30 : Représentation schématique de la validation croisée *leave-one-out*

Il est aussi possible de diviser le jeu de données en un nombre de sous-groupes définis, et utiliser cette méthode en retirant non pas un échantillon, mais un groupe d'échantillons (296).

Lorsque le nombre d'échantillons est assez important pour permettre une séparation en un groupe de calibration et de validation, il est conseillé d'utiliser des algorithmes de répartition, tel que l'algorithme de Kennard-Stone, afin de partitionner objectivement les échantillons et ainsi éviter le biais humain (318). Il peut être cependant pertinent de choisir subjectivement le groupe de validation, quand il permet de démontrer que la prédiction du modèle s'affranchit de variables inhérentes au jeu de données (279). Dans l'analyse du vin, cela se traduit par exemple en choisissant un cépage différent de la validation, une région viticole différente ou un millésime différent.

## 2. Paramètres descriptifs d'un modèle de prédiction

Plusieurs paramètres statistiques vont décrire un modèle, et ainsi permettre de discuter la calibration et la validation. L'erreur de prédiction par exemple, est définie comme l'erreur quadratique de validation croisée (RMSECV : *Root Mean Square Error of Cross Validation*) quand la validation croisée est utilisée, et comme l'erreur quadratique de prédiction (RMSEP : *Root Mean Square Error of Prediction*) quand une validation externe est utilisée (296,319). Ces valeurs sont très intéressantes car elles indiquent l'incertitude moyenne à laquelle sera confronté la prédiction d'échantillons futurs. Elles se retrouvent parfois exprimées en pourcentage, en divisant l'erreur par la valeur moyenne de la mesure (288). Ce sont des indicateurs clairs et efficaces pour discuter de la validité d'un modèle, car en représentant directement l'erreur de la prédiction, il est possible d'admettre ou non son acceptabilité. Cette valeur d'erreur sera aussi utilisée pour la détermination du nombre de variables latentes du modèle, choisies dans le but de minimiser cette incertitude moyenne (6).

L'erreur de calibration (RMSEC : *Root Mean Square Error of calibration*) est aussi fréquemment utilisée (10). Il faut cependant être prudent avec cette mesure car comme cela a été évoqué précédemment, une calibration performante n'implique pas forcément une prédiction performante. Néanmoins, cet indicateur peut s'avérer utile en comparaison avec l'erreur de prédiction. Lorsque les deux valeurs sont proches, la prédiction sera aussi performante que la calibration. Cela implique soit un modèle prédictif quand l'erreur est acceptable, soit une calibration qui n'explique pas correctement la variabilité du modèle quand l'erreur est trop élevée. A l'inverse, l'éloignement de ces deux valeurs traduit un modèle peu prédictif, qui n'est pas capable d'expliquer la variabilité de nouveaux échantillons en fonction de sa calibration.

Ces paramètres d'erreurs sont souvent associés au coefficient de régression linéaire ( $R^2$ ) qui permet de juger de la qualité de la régression. L'erreur standard de prédiction (SEP : *standard error of prediction*) est aussi reporté dans certaines publications en remplacement de l'erreur quadratique de prédiction (309,320,321).

Le dernier paramètre à prendre en compte pour expliquer un modèle de prédiction est l'écart prédictif résiduel (RPD : *residual predictive deviation*). Il se calcule en divisant la déviation standard de la variable étudiée par l'erreur quadratique moyenne de prédiction ou de validation croisée. Selon la littérature, un RPD inférieur à 1,4 indique un modèle de prédiction qui n'est pas fiable, et ne peut être utilisé que comme un indicateur. Lorsque le RPD est supérieur à 1,4 et inférieur à 2, le modèle

commence à être suffisamment fiable pour être utilisé pour la prédiction, avec une marge d'erreur importante. Lorsque le RPD est supérieur à 2, le modèle commence à être considéré comme bon, et lorsqu'il est supérieur à 3, il est considéré comme excellent (6,10,270).

### **3. Application de la spectroscopie couplée à la régression PLS pour l'analyse des polyphénols du raisin, du moût et du vin**

Comme cela a été évoqué précédemment, les spectroscopies infrarouge moyen et UV-visible semblent tout indiquées pour l'analyse des polyphénols. Cependant, la complexité de ces composés additionnée à la complexité de la matrice du vin et du raisin rendent l'exploitation des spectres bruts très difficile voire impossible. Dès lors, l'utilisation de l'analyse de données multivariées, et en particulier la régression PLS, apparaît comme la solution idéale pour extraire efficacement l'information spectrale des polyphénols pour la construction de modèle de prédiction. Certaines méthodes spectroscopiques comme l'infrarouge par réflexion totale atténuée ou réflexion diffuse permettent d'analyser des composés solides, et ne seront pas abordées, cette étude ne portant que sur l'analyse d'échantillons liquides (322).

#### ***a) Utilisation de la spectroscopie UV-visible pour l'analyse des polyphénols du raisin, du moût et du vin***

Plusieurs études ont été conduites afin de prédire la concentration des composés phénoliques du vin et du raisin grâce à la spectroscopie UV-visible. Avant de s'intéresser exclusivement à cette aire spectrale, il est à noter que plusieurs travaux ont aussi cherché à combiner l'UV-visible et le proche infrarouge pour prédire la composition phénolique des moûts en fermentation et des vins.

##### ***(1) L'UV-visible combiné au proche infrarouge***

La première étude, conduite par Cozzolino et al. en 2004, combine le visible au proche infrarouge (400 nm à 2500 nm) et cherche à prédire les concentrations en malvidine-3-*O*-glucoside, en pigments polymérisés et en tanins sur des vins en fermentation (14). Les résultats obtenus montrent des prédictions très satisfaisantes pour la malvidine-3-*O*-glucoside et les pigments polymérisés, avec des écarts prédictifs résiduels RPD respectivement de 3,5 et 3,2. Les résultats pour les tanins sont quant à eux bien plus faibles, avec un RPD de 1,8. Si la réponse spectrale dans le visible peut être attribuée aux anthocyanes, l'auteur indique que la majorité de la variabilité du proche infrarouge est

sûrement due à la concentration en sucre et en éthanol. Il est notamment compliqué de différencier, entre 2200 et 2300 nm, l'impact des polyphénols à côté de ces molécules. Il est dès lors difficile de s'assurer que les prédictions sont le fruit de toute la zone spectrale, et pas seulement du visible.

Une étude plus récente du même laboratoire a concentré ses recherches sur les mêmes longueurs d'ondes, mais cette fois ci sur le raisin et les teneurs en anthocyanes (323). Ces dernières sont extraites des baies de raisin grâce à une solution d'eau et d'éthanol 50/50 v/v, et ce travail soulève un point important quant à l'analyse sur du matériel solide : contrairement à des composés comme les acides ou les sucres, les polyphénols nécessitent une extraction préalable pour être étudiés. Les résultats montrent de nouveau une prédiction robuste, avec un  $R^2$  de validation croisée de 0,89. L'auteur a aussi montré que si l'on soustrayait la réponse spectrale de deux échantillons avec des concentrations en anthocyanes différentes, la majorité de l'absorbance restante se situait dans le visible. Ainsi, il semble confirmer que le proche infrarouge a un impact négligeable dans la mesure des anthocyanes.

La dernière étude portant sur l'analyse des polyphénols grâce au proche infrarouge combine aussi l'UV-visible, et balaye les longueurs d'onde entre 200 et 2500 nm (10). Les résultats avancés ne sont pas aussi performants que les précédents, avec un RPD pour la malvidine-3-*O*-glucoside, la catéchine et l'épicatéchine de respectivement 1,65, 1,66 et 1,17. Cependant, nous constatons un écart moins important entre les anthocyanes et les flavan-3-ols, qui pourrait indiquer que l'ajout de l'UV apporte des informations importantes sur ces derniers. Nous notons aussi plusieurs biais qui peuvent influencer les résultats. Tout d'abord, le nombre d'échantillons peut être considéré comme faible, avec seulement 39 vins différents. De plus, le prétraitement des spectres est relativement lourd (normalisation, première dérivée et correction de la ligne de base par exemple pour la catéchine), et diffère pour chaque composé. Le nombre de variables latentes sélectionnées n'est pas précisé, et si celui-ci est élevé, on peut être amené à penser que la prédiction est peut-être légèrement surestimée. Cette étude permet toutefois d'avancer que l'addition du proche infrarouge n'est pas forcément nécessaire, fait qui semble se confirmer avec les études précédentes.

## (2) *L'UV-visible*

La première étude répertoriée tentant de prédire la teneur en polyphénols dans les vins grâce à l'UV-visible couplé à la régression PLS date de 1995 (324). La marge d'erreurs obtenue sur un set de validation de 5 vins est de 11 % pour les tanins, et de 22 % pour les anthocyanes. Si le jeu d'échantillons est relativement faible (34 vins de Bordeaux), il permet déjà de mettre en évidence l'intérêt de cette méthode. L'auteur conclut que l'étude doit être approfondie avec un nombre plus

important de vins, mais que les résultats permettent d'avancer que la spectroscopie UV-visible est une candidate sérieuse pour une analyse rapide de routine des polyphénols.

En 2007, cette hypothèse est validée par les travaux de Skogerson et al. (2007). Les auteurs ont cherché à prédire les teneurs en anthocyanes, tanins et pigments polymérisés pendant différentes étapes de la fermentation (9). Le point fort de cette étude est l'utilisation d'un nombre important d'échantillons pour la calibration mais aussi pour la validation (200 pour chaque), ce qui permet d'écarter la possibilité d'une surestimation des prédictions. Les résultats avancés montrent des modèles robustes, avec pour les anthocyanes un  $R^2$  de prédiction de 0,89, et pour les tanins de 0,90. Cette précision décroît légèrement pour différencier les petits pigments polymérisés ( $R^2$  : 0,82) des grands pigments polymérisés ( $R^2$  : 0,63). Le seul biais notable de cette étude est l'utilisation d'un seul millésime, qui nécessiterait d'être étendue à plusieurs années pour valider une utilisation généralisée. Ces travaux comparent aussi les résultats obtenus avec d'autres en cours utilisant le proche infrarouge sur les mêmes échantillons, et valide l'intérêt limité de cette zone spectrale, avec des prédictions inférieures.

Deux études plus récentes rapportent des prédictions intéressantes pour les tanins précipitables avec une protéine ou un polysaccharide. L'avantage du dosage par précipitation, notamment avec une protéine, est que l'astringence apparaît comme fortement corrélée avec ce type de tanins (325). La première, conduite par Damberg en 2012 s'est focalisée sur la précipitation à la méthylcellulose corrélée à l'UV, et la faculté de la calibration à être transférée à un autre laboratoire (326). Ces travaux, effectués sur des vins finis commerciaux australiens, présentent des résultats de prédictions extrêmement robustes, avec un  $R^2$  de prédiction de 0,92 et un RPD de 3,5. Deux informations importantes sont alors à retenir : premièrement, la région de l'UV a un impact important dans la prédiction des tanins précipitables, et deuxièmement, le fait que la calibration soit transférable entre laboratoires permet d'envisager le développement de prédictions pour un usage généralisé.

La deuxième étude porte sur la comparaison du dosage à la méthylcellulose et du dosage à l'albumine de sérum bovin, en utilisant l'UV-visible couplé à la régression PLS (11). Si les deux méthodes utilisées se basent sur un principe similaire, des divergences sont notables. Tout d'abord, le fait d'utiliser un polysaccharide d'un côté et une protéine de l'autre change la nature des composés précipités (275,327). De plus, le dosage à la méthylcellulose précipite sans différencier pigments polymériques et tanins, là où le dosage à l'albumine de sérum bovin fait la distinction. Dès lors, ces deux analyses, bien que proches dans leur exécution, ne sont pas fortement corrélées (11). Cependant,

les résultats de prédiction avancés sont très similaires, avec un RPD de 2,97 pour l'albumine de sérum bovin, et de 3,5 pour la méthylcellulose. Ainsi, l'UV-visible présente un avantage certain pour étudier les différences dans la composition en tanins condensés.

L'étude la plus récente répertoriée a été conduite sur un large panel d'échantillons en fermentation et de vins d'Afrique du Sud, et a démontré la pertinence de la spectroscopie UV-visible couplée à la chimométrie pour prédire un grand nombre de composés phénoliques (12). Ces travaux valident la possibilité de prédire les concentrations en tanins précipitables ou en anthocyanes totales grâce à cette méthode. Cependant, pour des composés spécifiques comme les monomères d'anthocyanes, les données sont discutables. En effet, pour ces molécules, les résultats montrent soit des prédictions insuffisantes, soit la mise à l'écart de nombreux échantillons de la modélisation pour obtenir une calibration et une validation correctes.

De manière générale, par son habilité à détecter les molécules colorées, les doubles liaisons et les liaisons  $\pi$ , l'UV-visible couplé à la régression PLS présente un réel intérêt pour l'analyse des tanins condensés et des anthocyanes. Le fait d'éviter l'absorbance des principaux composés de la matrice du vin est à la fois un avantage et un inconvénient. Si cela permet de faciliter la prédiction, cette méthode ne pourra être adaptée qu'aux polyphénols, et ne permet pas de doser dans les mêmes conditions les autres composés du raisin, du moût et du vin.

### ***b) Utilisation de la spectroscopie infrarouge pour l'analyse des polyphénols du vin***

L'infrarouge à transformée de Fourier couplé à des outils de chimométrie présente aussi un réel intérêt pour le dosage des tanins et des anthocyanes. Cette aire spectrale contient bien plus d'informations que l'UV-visible, et son principal avantage est d'avoir été déjà développée pour la quantification de différents composés du vin, tels que les sucres, l'éthanol, ou différents acides (298–300). Ainsi, cette méthode a été testée pour quantifier les teneurs en polyphénols des moûts, vins et raisins.

La première étude répertoriée date de 2006, et tente de prédire la composition en tanins, anthocyanes et pigments polymérisés dans les vins (289). Si les résultats de prédiction sont très bons, avec une erreur de validation croisée majoritairement inférieure à 10 %, ils sont largement discutables. En effet, le set d'échantillons est très réduit (20 vins), et le nombre de variables latentes en comparaison très

élevé (10). Ainsi, il est difficile de juger de la pertinence de l'IRTF sur cette étude seule, car la prédiction n'est pas fiable, et peut conduire à des résultats largement surestimés.

Il faudra attendre l'année suivante pour mieux appréhender l'efficacité de cette technique. Les travaux de Fernández et Agosin ont cherché à déterminer la composition en tanins précipitables par les protéines, en tanins totaux et le degré de polymérisation moyen en utilisant l'infrarouge moyen (4000-650  $\text{cm}^{-1}$ ) (288). Les résultats sont très encourageants, avec un pourcentage d'erreur de prédiction aux alentours de 10 %. Néanmoins, les échantillons nécessitent d'être préalablement purifiés sur une colonne  $\text{C}_{18}$  SPE, ce qui ne permet pas d'adapter la méthode à un dosage rapide de routine.

Les premiers travaux qui montrent véritablement la pertinence de l'IRTF pour la prédiction des teneurs en anthocyanes et tanins des vins ont été produits par Soriano et al. en 2007 et par Jensen et al. en 2008. Ils ont été conduits tous les deux sur un spectromètre infrarouge moyen à transformée de Fourier Winescan FT-120 (5012- 926  $\text{cm}^{-1}$ ). La première étude se focalise sur la concentration en anthocyanes moléculaires et totales de 350 vins espagnols jeunes (13). Les modèles de prédiction pour les anthocyanes moléculaires sont assez hétérogènes, avec des  $R^2$  de prédiction variant de 0,64 à 0,94. Les plus robustes n'étant pas nécessairement les anthocyanes les plus concentrées, il est possible que certains résultats soient le fruit d'une corrélation indirecte. Il est par exemple étonnant de remarquer que la malvidine-3-*O*-acétylglucoside est mieux prédite que la malvidine-3-*O*-glucoside, alors qu'à l'inverse, la cyanidine-3-*O*-glucoside est mieux prédite que la cyanidine-3-*O*-acétylglucoside. Cependant, les valeurs de prédiction pour les anthocyanes totales sont très robustes ( $R^2$  : 0,92) et permettent de mettre en évidence le potentiel de l'IRTF. La seconde étude a mesuré la concentration en tanins par précipitation avec l'albumine de sérum bovin sur 128 vins finis commerciaux, et modélisé ces données de concentrations par régression PLS (15). Les résultats présentés sont très concluants, avec un  $R^2$  de validation de 0,94 après une sélection de longueurs d'ondes spécifiques.

Plus récemment, des prédictions robustes ont été présentées pour les phénols totaux, les flavonoïdes totaux et les anthocyanes totales sur des vins en fermentations, en se basant sur le proche infrarouge étendu et l'infrarouge moyen (328). Cependant, l'utilisation de lectures d'absorbance à 280 nm et 540 nm sans prétraitement pour calculer la concentration respective des flavonoïdes et des anthocyanes peut être sujet à discussion.

Les concentrations des composés phénoliques dans le raisin n'ont été que très peu étudiées, car comme nous l'avons évoqué précédemment, cela nécessite une extraction préalable. Des travaux se

sont pourtant focalisés sur la prédiction des anthocyanes du raisin, en utilisant un Winescan FT-120 (5012- 926  $\text{cm}^{-1}$ ) (329). Il a ainsi été montré qu'il était possible d'obtenir des prédictions efficaces pour les anthocyanes avec une extraction rapide par broyage. Les baies de raisin ont été mixées dans un broyeur pendant 30 secondes, puis filtrées sous vide à l'aide d'un papier filtre de 42  $\mu\text{m}$  de diamètre de pore maximum. Cependant, cette méthode ne semble pas adaptée aux tanins qui sont plus difficilement extractibles, et limite donc son utilisation.

L'une des études les plus récentes a recherché la différence entre le proche infrarouge, l'infrarouge moyen et l'infrarouge par réflexion totale atténuée pour prédire de nombreux composés phénoliques de vins en fermentation et vins finis commerciaux (287). Ainsi, la composition en anthocyanes, en tanins précipitables et tanins totaux a été approfondie. Les résultats avancés indiquent que la méthode la plus efficace est l'infrarouge moyen couplé à la régression PLS, et conduisent à des RPD de validation supérieure à 2 pour pratiquement tous les paramètres étudiés. Ces travaux valident donc l'infrarouge moyen comme l'une des méthodes les plus adaptées pour la quantification des molécules organiques, et ce malgré une réponse spectrale importante et parfois difficile à extraire.

Pour finir, malgré la complexité de la matrice du vin, l'utilisation de l'infrarouge moyen a aussi été reportée pour la prédiction des ellagitanins (330). Cette étude s'est concentrée sur la possibilité de prédire la concentration totale en ellagitanins de vins vieillissants avec différents apports de bois (chips de chêne, barriques de chêne français, barriques de chêne américain). Si les résultats énoncés semblent prometteurs, le manque d'échantillons utilisés et de robustesse de la validation ne permettent pas actuellement de prédire la concentration de ces composés dans le vin.

## **D. Conclusion générale**

Les polyphénols du raisin et du vin montrent une grande diversité et un impact important sur les propriétés organoleptiques du produit fini. Les anthocyanes et les tanins condensés, particulièrement, sont majoritairement responsables de la couleur, l'astringence et l'amertume des vins. Les compositions, concentrations et extractions des polyphénols sont largement influencées par les pratiques culturelles et œnologiques, et le vinificateur peut avoir un impact significatif sur leurs propriétés et distributions finales. Il apparaît dès lors crucial de pouvoir suivre leur évolution durant les différentes étapes de la vinification. Cependant, les analyses des polyphénols utilisées à ce jour demandent des ressources en matériels, en main d'œuvres, en temps, ainsi que des connaissances et des compétences qui ne les rendent pas adaptées à l'analyse de routine attendue par les praticiens.

Les spectroscopies UV-visible et IRTF apparaissent comme une alternative efficace dans le développement d'analyses rapides. Ces deux aires spectrales permettent de caractériser et doser les polyphénols, et particulièrement les anthocyanes et les tanins. Néanmoins, la grande variabilité de la matrice du raisin, du moût et du vin ne permet pas une lecture spécifique avec un traitement direct des spectres. La solution la plus adaptée est donc d'appliquer des outils chimiométriques et plus précisément la régression PLS à ces données, afin d'extraire l'information essentielle.

Plusieurs études ont démontré l'intérêt de la régression PLS couplée à la spectroscopie IRTF ou UV-visible. Elles ont mis en lumière les facultés de cette méthode dans le développement de modèles de prédiction, mais au regard de la diversité des échantillons, des analyses de références, des prétraitements des spectres et de la régression multivariée, il est compliqué d'obtenir une appréciation globale de la robustesse de ces modèles et de les comparer entre eux. Il est ainsi difficile de juger de la reproductivité des résultats et de leur application en milieu industriel. A ce jour, aucun travail n'a permis de comparer la spectroscopie UV-visible et l'IRTF, ou encore la combinaison possible de deux techniques spectroscopiques pour tenter d'accroître la précision des modèles.

---

## **II. Matériels et méthodes**

---



## A. Matériels

### 1. Réactifs, solvants et standards

Tous les produits utilisés pour l'étude sont de qualités analytiques et répertoriés dans le Tableau 7. Trois fournisseurs ont été utilisés : Sigma-Aldrich (Saint Quentin Fallavier, France), Fisher Scientific (Geel, Belgique) et VWR-Prolabo (Fontenay-sous-Bois, France).

Produit	Pureté	Fournisseur
<b>Solvants</b>		
Acétonitrile	> 99,5 %	VWR Prolabo
Éthanol	> 99,9 %	VWR Prolabo
Méthanol	100 %	VWR Prolabo
<b>Acides</b>		
Acide acétique glacial	100 %	Fisher Scientific
Acide L-ascorbique	> 99 %	Sigma-Aldrich
Acide chlorhydrique	37 %	Fisher Scientific
Acide formique	> 99 %	Fisher Scientific
Acide L-tartrique	> 99,5 %	Sigma-Aldrich
<b>Bases</b>		
Hydroxyde de sodium	> 98 %	Sigma-Aldrich
Triéthylamine	> 99,5 %	Sigma-Aldrich
<b>Réactifs</b>		
Bisulfite de sodium	40 %	Sigma-Aldrich
Chlorure de fer III hydrate (6 H <sub>2</sub> O)	> 97 %	Sigma-Aldrich
Chlorure de sodium	> 99 %	Sigma-Aldrich
Dodécylsulfate de sodium	> 99 %	Sigma-Aldrich
Métabisulfite de sodium	> 99 %	Sigma-Aldrich
Méthylcellulose (viscosité : 1400 cP)	100 %	Sigma-Aldrich
Phloroglucinol	> 99 %	Sigma-Aldrich
Albumine de sérum bovin	> 96 %	Sigma-Aldrich
Sulfate d'ammonium	> 99 %	Sigma-Aldrich
<b>Standards</b>		
(+)-catéchine hydrate (1 H <sub>2</sub> O)	> 98 %	Sigma-Aldrich
(-)-épicatéchine	> 90 %	Sigma-Aldrich
Malvidine-3-O-glucoside chloride	> 95 %	Sigma-Aldrich

Tableau 7 : Solvants, acides, bases, réactifs et standards utilisés

L'eau ultra pure utilisée pour l'étude a été purifiée via un système Milli-Q (Millipore, Bedford, MA, USA). L'eau ainsi obtenue a une résistivité de 18,2 M $\Omega$ •cm (à 25 °C) et une valeur de carbone organique total inférieure à 5 ppb.

## **2. Echantillonnage**

L'échantillonnage est réalisé en favorisant la diversité et la variabilité, afin de concevoir des modèles de prédiction représentatifs de l'évolution du raisin et vin.

### ***a) Echantillons de raisin***

L'échantillonnage de raisins est réalisé avec l'Union de Producteurs de Saint-Emilion (UdPSE). Cette cave coopérative gère 700 hectares de vignobles, et propose une grande variabilité de raisins et de vins, issus de terroirs et de pratiques culturelles différents. Tous les échantillons prélevés appartiennent à des parcelles et cépages destinés à la vinification de vins commerciaux. La sélection des raisins est menée en recherchant la variabilité de la qualité phénolique et technologique des raisins.

Pour chaque échantillon, un protocole est établi. Le prélèvement est effectué le matin, et 1 kg de raisins est récolté. Les baies sont extraites sur différentes zones de la grappe et de la vigne sur toute la longueur d'un rang de la parcelle sélectionnée. La sélection se veut homogène et représentative de la parcelle.

Millésime 2017 : 6 parcelles ont été suivies pendant quatre semaines avant les vendanges. 4 parcelles de Merlot et 2 parcelles de Cabernet franc sont sélectionnées pour refléter la diversité proposée par l'UdPSE. Un échantillonnage par semaine est effectué pour suivre la maturité, pour un total de 24 échantillons.

Millésime 2018 : 6 parcelles différentes, 4 de Merlot et 2 de Cabernet franc sont suivies pendant quatre semaines avant les vendanges. Un échantillon est prélevé chaque semaine soit un total de 24 échantillons.

Millésime 2019 : 15 nouvelles parcelles de Merlot sont suivies, avec un prélèvement par semaine pendant cinq semaines jusqu'aux vendanges (75 échantillons).

### ***b) Echantillons de vins en fermentation alcoolique***

L'échantillonnage de vins pendant la fermentation alcoolique est réalisé à l'UdPSE. Le schéma de vinification est identique pour tous les vins. Le raisin, égrappé et foulé, est fermenté dans des cuves en acier inoxydable. Il subit une première période de macération à froid à 4 °C pendant 4 jours, suivie d'une période de fermentation alcoolique d'une à deux semaines en présence de levures commerciales (*Saccharomyces cerevisiae*), puis une macération post-fermentaire d'une à deux semaines jusqu'à l'écoulage. Tous les vins suivis sont destinés à être commercialisés.

Chaque échantillon est prélevé le matin après un remontage, au moment où la cuve est la plus homogène possible. Une demi-bouteille (37,5 cL) est récupérée, sulfitée à 2 g/hL, et bouchée à l'aide d'un obturateur synthétique. L'échantillon est alors congelé à -22 °C jusqu'à son analyse. Tous les échantillons ont été traités dans les six mois qui ont suivi le prélèvement.

Millésime 2017 : 6 cuves (4 de Merlot et 2 de Cabernet franc) ont été suivies pendant la période de fermentation alcoolique. Un échantillon par cuve est prélevé tous les deux jours, du premier jour de vinification jusqu'à l'écoulage. Au total, 105 échantillons sont disponibles ce qui représente 15 à 20 échantillons par cuve.

Millésime 2018 : 6 cuves (4 de Merlot et 2 de Cabernet franc) sont suivies. Un échantillon par cuve est prélevé tous les deux jours pendant la période de macération post-fermentaire et pendant la fermentation alcoolique. Puis un échantillon est prélevé tous les 3-4 jours pendant la macération post-fermentaire jusqu'à l'écoulage. Au total, 65 échantillons sont collectés (10 échantillons par cuve). Il s'agit de la fermentation des parcelles qui ont été suivies lors de la maturité du raisin la même année.

Millésime 2019 : 9 cuves (7 de Merlot et 2 de Cabernet franc) sont suivies. Le même protocole d'échantillonnage que pour 2018 a été appliqué. Un total de 90 échantillons sont disponibles (10 échantillons par cuve).

### ***c) Echantillons de vins durant le vieillissement en barriques***

Afin d'évaluer l'incidence du vieillissement en barriques et de comparer à un vieillissement en bouteilles, 4 vins de Merlot ont été suivis durant le millésime 2018. Ces échantillons sont récupérés à l'UdPSE. Les vins ont été transférés en barriques neuves de chêne français à la fin de la fermentation malolactique, et sulfités à 2 g/hL. Un premier prélèvement a eu lieu lors du

transfert, à raison de 5 bouteilles de 75 cL par barrique suivie. Un prélèvement est alors effectué tous les deux mois pendant 8 mois, et les échantillons barriques et bouteilles sont analysés en parallèle.

#### *d) Echantillons de vins finis commerciaux*

92 vins sont sélectionnés pour l'étude, afin de préparer des modèles de prédiction reflétant une diversité d'origines et de types de vins. Le millésime, l'origine et les cépages de tous les vins étudiés sont reportés dans le Tableau 8. Les vins surlignés en bleu sont ceux qui ont été testés par le panel de dégustateurs entraînés.

	<b>Millésime</b>	<b>Origine</b>	<b>Cépage majoritaire</b>	<b>Cépage minoritaire</b>
<b>Vin 1</b>	2015	France	Merlot	Cabernet franc
<b>Vin 2</b>	2015	France	Merlot	
<b>Vin 3</b>	2015	France	Merlot	Cabernet franc
<b>Vin 4</b>	2016	France	Cabernet sauvignon	
<b>Vin 5</b>	2017	France	Cabernet sauvignon	
<b>Vin 6</b>	2017	France	Cabernet franc	
<b>Vin 7</b>	2017	France	Cabernet franc	
<b>Vin 8</b>	2017	France	Tannat	
<b>Vin 9</b>	2015	France	Merlot	
<b>Vin 10</b>	2016	France	Merlot	
<b>Vin 11</b>	2017	France	Merlot	
<b>Vin 12</b>	2017	France	Cabernet sauvignon	
<b>Vin 13</b>	2017	France	Cabernet franc	
<b>Vin 14</b>	2015	France	Cabernet sauvignon	
<b>Vin 15</b>	2016	France	Cabernet sauvignon	
<b>Vin 16</b>	2017	France	Cabernet sauvignon	
<b>Vin 17</b>	2015	France	Merlot	
<b>Vin 18</b>	2016	France	Merlot	
<b>Vin 19</b>	2014	France	Tannat	
<b>Vin 20</b>	2014	France	Syrah	
<b>Vin 21</b>	2016	France	Syrah	
<b>Vin 22</b>	2015	France	Pinot noir	
<b>Vin 23</b>	2015	France	Pinot noir	
<b>Vin 24</b>	2015	France	Pinot noir	
<b>Vin 25</b>	2014	France	Merlot	
<b>Vin 26</b>	2015	France	Tempranillo	
<b>Vin 27</b>	2017	France	Merlot	
<b>Vin 28</b>	2016	France	Pinot noir	
<b>Vin 29</b>	2017	France	Pinot noir	
<b>Vin 30</b>	2016	France	Carignan	
<b>Vin 31</b>	2017	France	Carignan	
<b>Vin 32</b>	2017	France	Cabernet sauvignon	
<b>Vin 33</b>	2017	France	Merlot	
<b>Vin 34</b>	2015	France	Cinsault	
<b>Vin 35</b>	2015	France	Syrah	
<b>Vin 36</b>	2014	France	Syrah	
<b>Vin 37</b>	2017	Argentine	Malbec	
<b>Vin 38</b>	2016	France	Gamay	
<b>Vin 39</b>	2017	France	Gamay	
<b>Vin 40</b>	2015	France	Merlot	
<b>Vin 41</b>	2015	France	Merlot	Cabernet sauvignon
<b>Vin 42</b>	2014	France	Merlot	Cabernet sauvignon, Petit Verdot

Vin 43	2015	France	Cabernet sauvignon	Merlot, Cabernet franc, Petit Verdot
Vin 44	2014	France	Cabernet sauvignon	Merlot, Malbec
Vin 45	2006	France	Cabernet sauvignon	Merlot, Petit Verdot
Vin 46	2016	France	Niellucciu	Sciaccarellu
Vin 47	2014	Espagne	Tempranillo	
Vin 48	2017	Grèce	Agiorgitiko	Syrah
Vin 49	2015	Italie	Sagniovese	
Vin 50	2015	France	Syrah	Mourvèdre, Grenache
Vin 51	2017	France	Cabernet franc	
Vin 52	2016	France	Cabernet franc	Cot, Pinot noir
Vin 53	2014	France	Gamay	Cot, Pinot noir
Vin 54	2014	Etats-Unis	Petite syrah	
Vin 55	2015	Portugal	Jaen	Touriga Nacional
Vin 56	2015	France	Syrah	Cabernet sauvignon, Carignan, Grenache
Vin 57	2016	France	Syrah	
Vin 58	2015	France	Grenache	Carignan, Syrah
Vin 59	2015	France	Syrah	Grenache, Mourvèdre
Vin 60	2016	France	Grenache	Syrah, Mourvèdre
Vin 61	2017	France	Grenache	Syrah, Carignan
Vin 62	2010	France	Mondeuse	
Vin 63	2016	France	Merlot	Malbec
Vin 64	2017	France	Malbec	
Vin 65	2016	France	Negrette	
Vin 66	2015	France	Mansois	Syrah
Vin 67	2016	France	Mansois	
Vin 68	1997	Allemagne	Trollinger	
Vin 69	2002	Italie	Nero d'Avola	
Vin 70	2016	Argentine	Malbec	
Vin 71	2005	Argentine	Malbec	
Vin 72	2010	Argentine	Malbec	
Vin 73	2015	Italie	Nero d'Avola	
Vin 74	2007	Etats-Unis	Zinfandel	
Vin 75	2015	Italie	Zinfandel	
Vin 76	2015	Autriche	Zweigelt	
Vin 77	2017	Afrique du Sud	Grenache	
Vin 78	2012	Italie	Negroamaro	
Vin 79	2011	Chili	Carmenere	
Vin 80	2014	Chili	Carmenere	
Vin 81	2013	Italie	Barbera	
Vin 82	2001	Chili	Carmenere	
Vin 83	2010	Danemark	Rondo	
Vin 84	2012	Australie	Petite syrah	
Vin 85	2013	Afrique du Sud	Petit verdot	
Vin 86	2003	Espagne	Petit verdot	
Vin 87	2005	Chili	Petite syrah	
Vin 88	2016	Afrique du Sud	Pinotage	
Vin 89	2016	Afrique du Sud	Pinotage	
Vin 90	2002	Australie	Petite syrah	
Vin 91	2013	Italie	Nebbiolo	
Vin 92	2000	Italie	Sagniovese	

Tableau 8 : Millésime, origine et cépages des vins finis commerciaux de l'étude (les vins surlignés ont été testés par le panel de dégustateurs entraînés)

La répartition des millésimes des vins est présentée en Figure 31, ainsi que la répartition des cépages principaux (Figure 32). Sur les 92 vins, 72 sont monocépages alors que 20 sont issus d'assemblages.

La variabilité de la provenance géographique est aussi prise en compte. Sur l'ensemble des échantillons, 61 sont originaires de France, pour représenter la diversité des vignobles du pays. Pour le reste des vins, 8 proviennent d'Italie, 4 d'Argentine, 4 du Chili et 4 d'Afrique du Sud, 2 d'Australie, 2 d'Espagne et 2 des Etats Unis, et 1 d'Autriche, 1 du Danemark, 1 de la Grèce et 1 du Portugal.

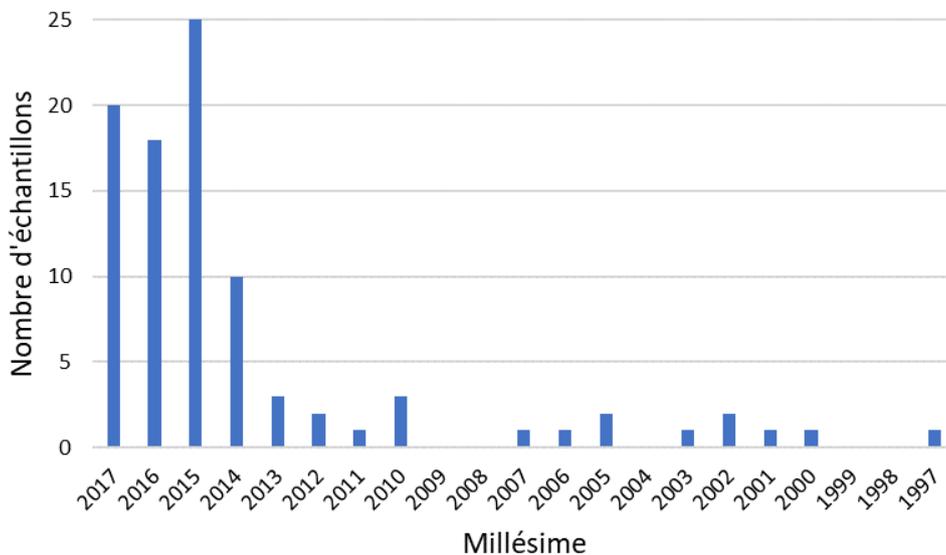


Figure 31 : Répartition des millésimes du panel d'échantillons de vins finis commerciaux

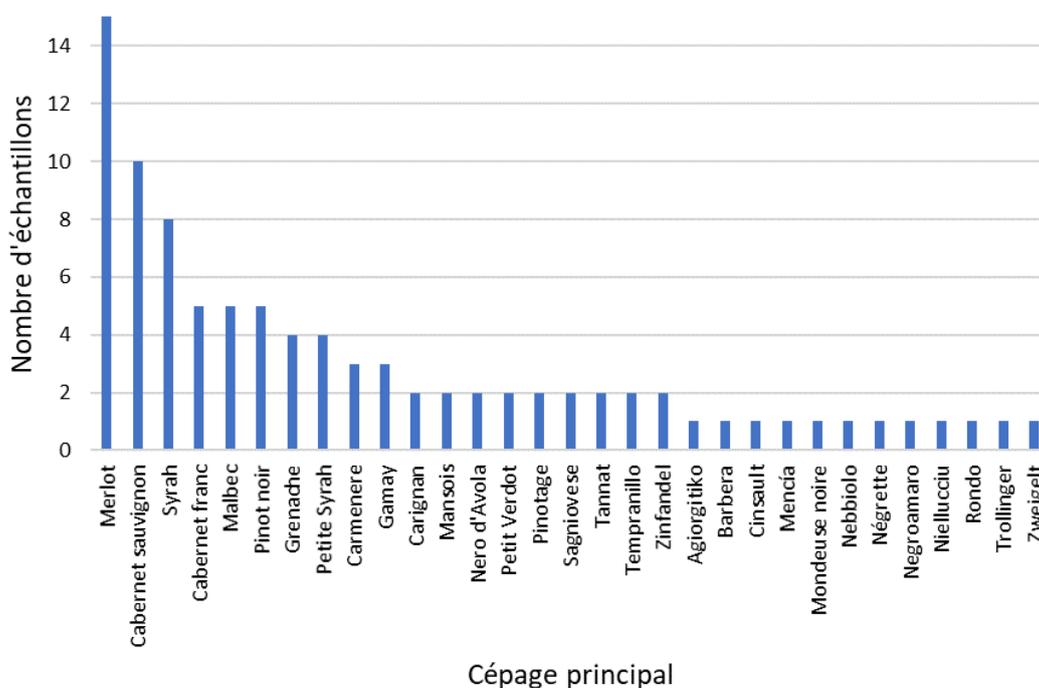


Figure 32 : Répartition des cépages principaux du panel d'échantillons de vins finis commerciaux

### ***e) Prétraitement des échantillons avant analyses***

Afin d'obtenir un échantillon utilisable en IRTF et UV-visible, certains prétraitements sont nécessaires. Il doit être limpide pour éviter les biais d'absorbances, et refléter la composition phénolique du produit.

Pour les échantillons de raisins, une extraction préalable par une solution à pH 3,2 suivie d'une filtration a été effectuée. Elle sera détaillée ultérieurement. Il s'agit d'une extraction longue, mais il a été décidé de se focaliser en priorité sur une méthode d'extraction connue et détaillée dans la littérature pour juger de l'efficacité de l'IRTF et l'UV-visible dans la prédiction des teneurs en tanins et anthocyanes des raisins (331,332).

Les échantillons de vins en fermentation ont été décongelés puis centrifugés à 10000 g pendant 5 min à température ambiante, et seul le surnageant a été conservé pour les analyses.

Les échantillons pendant l'élevage des vins et les vins finis commerciaux n'ont subi aucun prétraitement.

Tous les échantillons ont été conservés à 4 °C pendant le temps des analyses, dans des récipients fermés sous atmosphère d'azote pour éviter toute oxydation.

## **B. Analyses spectroscopiques**

Les analyses spectroscopiques ont été réalisées en triplicat. Pour chaque échantillon, un spectre infrarouge à transformée de Fourier et un spectre Ultraviolet-Visible ont été enregistrés.

### **1. Mesure des spectres UV-visible**

La mesure des spectres UV-visible est effectuée sur un spectrophotomètre Jasco V-630 (JASCO, Japon), et pilotée par le logiciel Spectra Manager. L'instrument est muni d'une lampe deutérium pour la mesure UV et d'une lampe halogène pour la mesure visible. Les échantillons sont scannés après une dilution au 1/100 dans de l'eau ajustée à pH 3,3 avec de l'acide chlorhydrique, dans une cuvette en quartz de 10 mm de parcours optique avec l'eau comme blanc. Les spectres sont enregistrés en absorbance entre 200 et 700 nm, avec 1 nm d'intervalle entre chaque mesure. La zone a ensuite été redéfinie pour correspondre à la limite d'absorbance de l'instrument utilisé. Ainsi, il a été décidé que pour les spectres UV-visible, les absorbances supérieures à 1,2, et donc supérieures à la limite de détection du spectrophotomètre, ne seraient pas prises en compte. Le spectre obtenu est

compris entre 250 et 700 nm. Lorsque certaines références démontrent des zones d'intérêts pour les polyphénols dans le cas de valeurs d'absorbance inférieures à 250 nm, il a été décidé de ne pas diluer davantage l'échantillon afin de garder la partie visible de la réponse spectrale suffisamment importante pour l'analyse de la concentration des anthocyanes et des tanins via une seule lecture d'échantillons (325).

La Figure 33 présente les spectres d'absorbance pour le panel de vins finis commerciaux, afin d'illustrer la réponse spectrale observée pour les longueurs d'ondes sélectionnées.

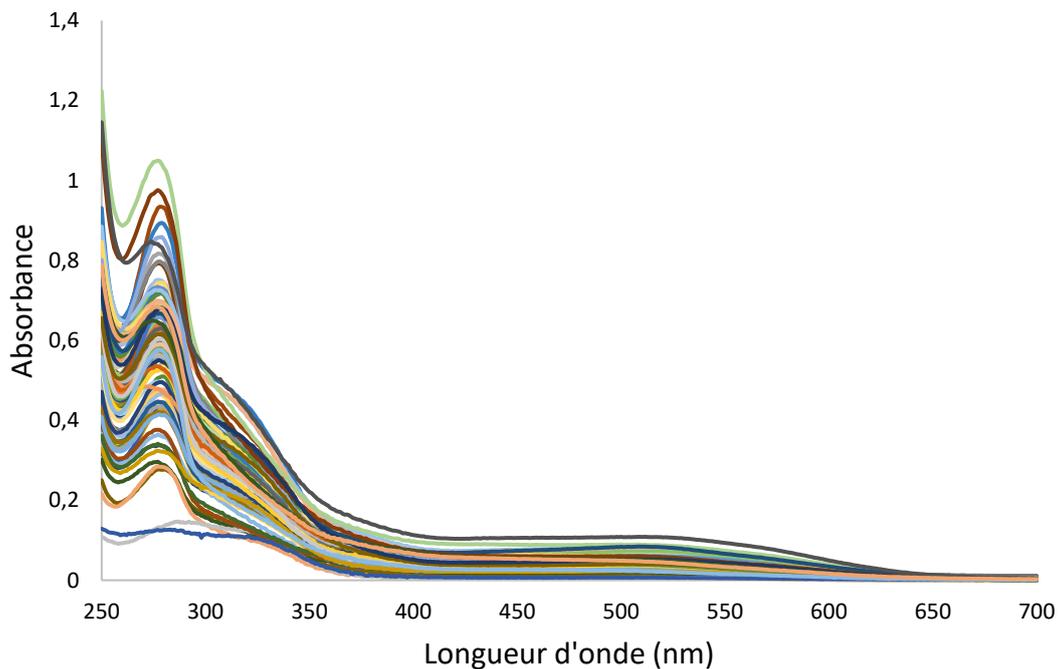


Figure 33 : Spectres d'absorbance UV-visible du panel de vins finis commerciaux

## 2. Mesure des spectres IRTF

La mesure des spectres IRTF est réalisée sur un WineScan Flex (FOSS, Hillerød, Danemark) équipé d'un interféromètre de Michelson. Un volume de 7 mL est pompé en deux fois, tout d'abord pour nettoyer le circuit, puis ensuite pour l'analyse, et passe dans une cuvette (37  $\mu\text{m}$  de parcours optique) thermostatée à 40 °C. Le circuit est nettoyé toutes les cinq analyses via une solution de nettoyage diluée dans l'eau (FOSS Cleaning solution, Hillerød, Danemark). Avant chaque utilisation et toutes les heures durant le fonctionnement, un zéro (eau) est mesuré pour s'assurer que la ligne de base ne dévie pas.

Le spectre enregistré est compris entre 925 et 5011  $\text{cm}^{-1}$  avec un intervalle de 3.858  $\text{cm}^{-1}$  entre chaque mesure. Il est enregistré en transmittance et transformé en absorbance via la formule suivante :

$$\text{Absorbance} = -\log(\text{Transmittance})$$

Tout comme pour les spectres UV-visible, des zones du spectre IRTF ont été éliminées pour ne garder que les longueurs d'ondes d'intérêt et éviter de mauvaises corrélations induites par du bruit de fond ou des réponses spectrales non assimilées aux polyphénols (15). Ainsi, les longueurs d'ondes suivantes ont été éliminées :

- Les longueurs d'ondes entre 925 et 1060  $\text{cm}^{-1}$ , qui correspondent à l'une des bandes d'absorption de l'eau. L'absorbance est alors trop saturée et ne permet pas d'extraire d'informations utiles.
- Les longueurs d'ondes entre 1543 et 1736  $\text{cm}^{-1}$ , qui correspondent à une zone de bruits de fond sans informations extractibles.
- Les longueurs d'ondes entre 1929 et 2507  $\text{cm}^{-1}$ , qui correspondent à une zone sans informations.
- Les longueurs d'ondes entre 3010 et 5011  $\text{cm}^{-1}$ , qui correspondent à une zone de bruits de fond très importante, ainsi qu'une zone sans informations pour les polyphénols.

La Figure 34 montre les spectres d'absorbance complets et réduits pour le panel de vins finis commerciaux. Nous constatons par comparaison que les longueurs d'ondes éliminées permettent de réduire considérablement le bruit de fond et les zones sans informations. Cette sélection apparaît être une première étape pour faciliter la conception de modèles de prédiction.

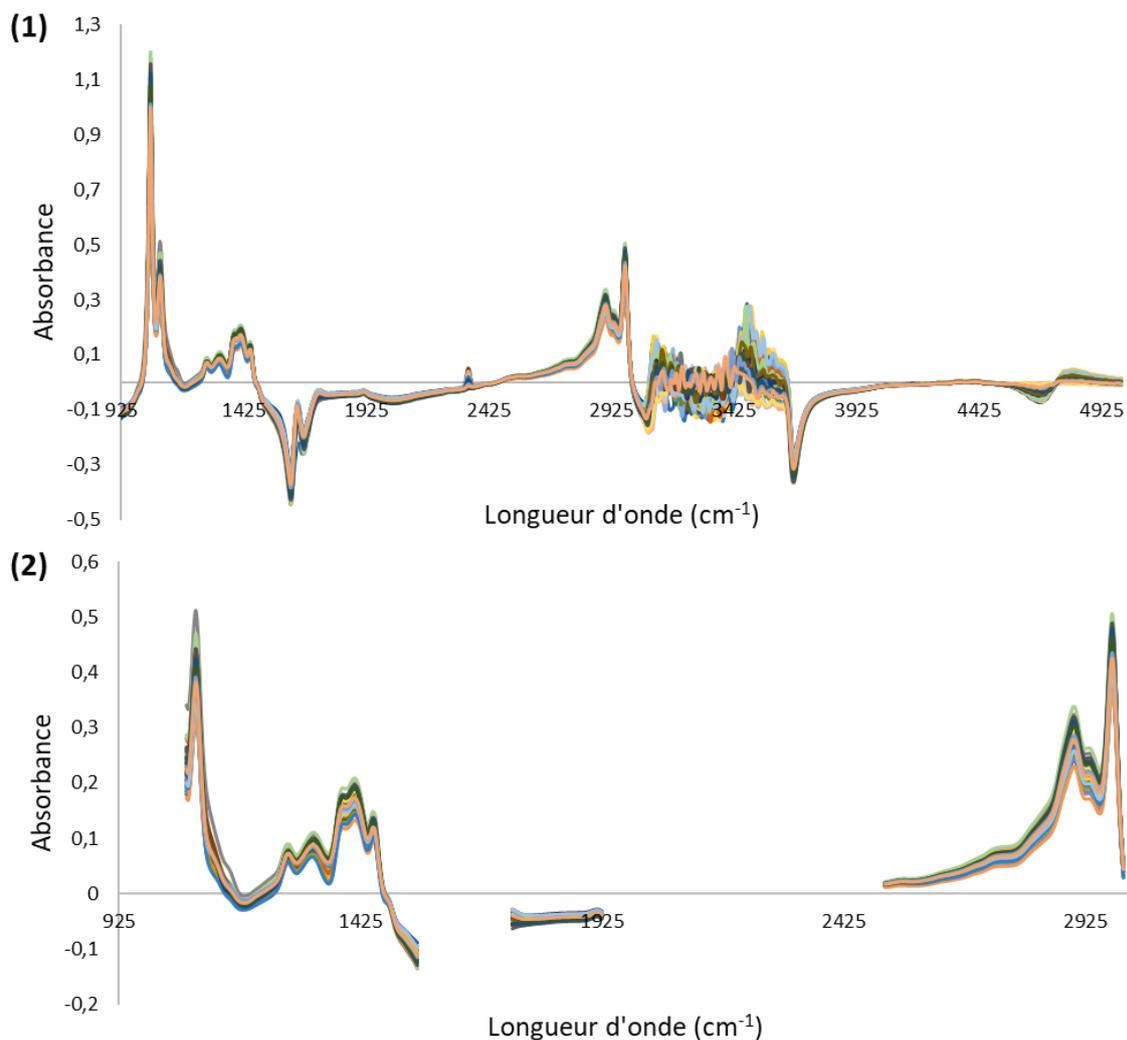


Figure 34 : (1) Spectres IRTF complets du panel de vins finis commerciaux, (2) Spectres IRTF avec les longueurs d'ondes sélectionnées du panel de vins finis commerciaux

### C. Analyses de références des anthocyanes et des tanins

Le choix des analyses de références pour la conception de modèles de prédiction est primordial. Celles-ci doivent être représentatives de la concentration globale des tanins et anthocyanes du milieu, et permettre de les caractériser précisément pour pouvoir suivre leur évolution.

Pour les anthocyanes, il est donc intéressant de pouvoir caractériser les composés sous forme monomérique ou libre, car c'est un bon indicateur de l'extraction pendant la fermentation alcoolique. La concentration en anthocyanes peut aussi agir comme un marqueur de la stabilité de la couleur, ou révéler des défauts comme une oxydation importante (333,334). De plus, afin de pouvoir évaluer

l'évolution de ces composés, un suivi du ratio entre pigments monomériques et pigments polymérisés s'impose. Il permettra de juger de la proportion d'anthocyanes qui se polymérise, et par extension de la proportion qui se dégrade ou précipite.

Concernant les tanins condensés, il est important de doser les teneurs en tanins totaux et précipitables, ces derniers montrant une réactivité accrue avec les protéines de la salive. Par ailleurs, disposer de spectres pour une gamme de concentrations de chaque monomère de flavan-3-ols apparaît aussi d'intérêt pour disposer de références. Certains de ces composés sont des marqueurs des tanins de la pellicule ou du pépin, et permettront de juger du déroulement de l'extraction lors de la vinification. Pour finir, le degré de polymérisation moyen impacte aussi les perceptions d'astringence, et peut s'avérer utile.

Afin de s'assurer de la répétabilité et de l'erreur due aux analyses de références, toutes les analyses ont été effectuées en triplicat.

## **1. Maturité phénolique déterminée par la méthode Glories**

Deux maturités sont importantes lors du développement de la baie de raisin : la maturité technologique déterminée par les teneurs en sucres et acides, et la maturité phénolique qui est liée à l'accumulation des tanins condensés et anthocyanes. Pour juger de cette dernière, Glories a développé une méthode d'analyse qui permet de déterminer le potentiel en anthocyanes extractibles et totales, la richesse polyphénolique totale ainsi que la maturité des pépins (331).

1 kg de baies de raisin sont broyées dans un blender à vitesse maximale pendant 1 min. 50 mL de ce broyat sont placés dans deux erlenmeyers de 250 mL. Ce volume est déterminé par pesée pour plus de précision. La masse en gramme est mesurée par la formule suivante,  $50 \times \text{densité}$ , où la densité est obtenue par lecture sur un densimètre optique. Dans le premier erlenmeyer sont ajoutés 50 mL d'une solution à pH 1 (Eau ajustée à pH 1 avec de l'acide chlorhydrique). Dans le second erlenmeyer sont ajoutés 50 mL d'une solution à pH 3,2 (5 g/L d'acide tartrique dilué dans l'eau, ajusté à pH 3,2 avec de l'hydroxyde de sodium). Après homogénéisation, les erlenmeyers sont bouchés au parafilm et laissés à macérer à température ambiante pendant 4 heures.

A la fin de la macération, les deux solutions sont filtrées sur laine de verre pour obtenir un filtrat limpide (Figure 35). A noter que la solution à pH 3,2, qui représente les polyphénols extractibles du raisin, sera utilisée comme solution pour les autres analyses de références. Il n'est pas possible

d'utiliser la solution à pH 1 en parallèle car celle-ci est trop acide pour l'instrument de spectroscopie IRTF, et elle risquerait d'endommager durablement l'appareillage.



Figure 35 : Etape de filtration de la méthode Glories, obtention des filtrats qui permettront l'analyse de la composition phénolique

La manipulation suivante est réalisée sur les deux solutions pH 1 et pH 3,2 en parallèle. 250  $\mu$ L de filtrat sont mélangés à 250  $\mu$ L d'éthanol acidifié avec 0,1 % d'acide chlorhydrique v/v, et 5 mL d'eau acidifiée avec 2 % d'acide chlorhydrique v/v. Cette solution mère est homogénéisée et après 10 min elle est répartie dans deux tubes :

- Dans le premier tube (réaction), 400  $\mu$ L de la solution mère sont mélangés à 1 mL d'une solution à 15 % de bisulfite de sodium dans l'eau v/v.
- Dans le second tube (témoin), 400  $\mu$ L de la solution mère sont mélangés à 1 mL d'eau.

Après 20 min, la différence d'absorbance à 520 nm est mesurée (10 mm de parcours optique avec l'eau distillée comme témoin). Pour finir, l'absorbance des polyphénols totaux est mesurée à 280 nm (10 mm de parcours optique avec l'eau distillée comme témoin dans une cuvette en quartz) en diluant 50 fois 1 mL du filtrat à pH 3,2.

Dès lors, il est possible d'obtenir le potentiel en anthocyanes extractibles ( $A_{pH3,2}$ ), le potentiel total en anthocyanes ( $A_{pH1}$ ), la richesse polyphénolique totale (RPT), les pourcentages d'anthocyanes extractibles (PAE%) et la maturité des pépins (MP) en suivant les formules suivantes :

$$A_{pH3,2} \text{ (mg/L)} = 2 \times 875 \times (A_{520_{pH3,2} \text{ témoin}} - A_{520_{pH3,2} \text{ réaction}})$$

$$A_{pH1} \text{ (mg/L)} = 2 \times 875 \times (A_{520_{pH1} \text{ témoin}} - A_{520_{pH1} \text{ réaction}})$$

$$RPT = 2 \times A_{280} \times 100$$

$$PAE \% = \frac{ApH_{3,2}}{ApH_1} \times 100$$

$$MP = \frac{RPT - ApH_{3,2} \times 0,04}{RPT} \times 100$$

## 2. Dosage des anthocyanes libres par décoloration au bisulfite de sodium

La concentration en anthocyanes libres dans les différents échantillons étudiés a été obtenue en utilisant leur propriété à se décolorer en présence de bisulfite de sodium (274). Un des principaux avantages de ce dosage est qu'il est utilisé comme une analyse de référence depuis plus de 50 ans et est bien connu des œnologues.

Le mode opératoire est identique à celui utilisé précédemment pour obtenir ApH1 et ApH3,2 à partir du filtrat. La concentration en anthocyanes libres est alors calculée grâce à la courbe d'étalonnage établie par Ribéreau-Gayon et Stonestreet et exprimée en équivalent malvidine-3-*O*-glucoside, suivant la formule suivante :

$$[C_{anthocyanes\ libres}](mg/L) = 875 \times (A_{520\text{témoin}} - A_{520\text{réaction}})$$

## 3. Quantification et caractérisation des anthocyanes moléculaires par chromatographie liquide haute performance (CLHP) et détection UV-visible

Douze anthocyanes monomères ont été caractérisées et quantifiées dans les échantillons par CLHP-UV-visible :

- Cinq anthocyanes monoglucosylées : la delphinidine-3-*O*-glucoside (Dp-3-*O*-glc), la cyanidine-3-*O*-glucoside (Cy-3-*O*-glc), la pétunidine-3-*O*-glucoside (Pt-3-*O*-glc), la péonidine-3-*O*-glucoside (Pn-3-*O*-glc) et la malvidine-3-*O*-glucoside (Mlv-3-*O*-glc).
- Cinq anthocyanes monoglucosylées acétylées : la delphinidine-3-*O*-acétylglucoside (Dp-3-*O*-acglc), la cyanidine-3-*O*-acétylglucoside (Cy-3-*O*-acglc), la pétunidine-3-*O*-acétylglucoside (Pt-3-*O*-acglc), la péonidine-3-*O*-acétylglucoside (Pn-3-*O*-acglc) et la malvidine-3-*O*-acétylglucoside (Mlv-3-*O*-acglc).
- Deux anthocyanes monoglucosylées coumaroylées : la péonidine-3-*O*-coumaroylglucoside (Pn-3-*O*-cmglc) et la malvidine-3-*O*-coumaroylglucoside (Mlv-3-*O*-cmglc).

Les échantillons sont injectés directement après filtration (diamètre de pores de 0,45 µm). L'analyse est réalisée sur une CLHP Thermo-Scientific *Accela series* (Thermo Fisher Scientific, Waltham, MA, USA), équipée d'une pompe *Accela 600*, d'un passeur d'échantillons et d'un détecteur à barrettes de diode (DAD). La longueur d'onde de détection utilisée est de 520 nm. La colonne utilisée pour la séparation des anthocyanes est une phase inverse Nucleosil C<sub>18</sub> (250 mm × 4 mm, 5 µm) (Agilent, Santa Clara, CA, USA). Le débit d'injection est réglé sur 1 mL/min, avec un volume d'injection de 20 µL à 15 °C. Les solvants utilisés sont (A) eau-acide formique (95/5, v/v), et (B) acétonitrile-acide formique (95/5, v/v), avec le gradient d'élution suivant (Tableau 9) :

<b>Temps (min)</b>	<b>0</b>	<b>25</b>	<b>35</b>	<b>40</b>	<b>41</b>	<b>46</b>
Solvant A (%)	90	65	0	0	90	90
Solvant B (%)	10	35	100	100	10	10

Tableau 9 : Gradient d'élution pour la séparation des anthocyanes moléculaires en CLHP

Afin de calculer les concentrations obtenues, une courbe de calibration a été préalablement établie, en utilisant la malvidine-3-*O*-glucoside. La gamme utilisée comporte 5 concentrations différentes (5, 10, 20, 50, 100, 200 mg/L) et a été injectée en triplicat. Les aires des pics obtenus sont donc reportées sur cette courbe pour obtenir un résultat exprimé en mg/L d'équivalent de malvidine-3-*O*-glucoside.

#### **4. Détermination du ratio pigments monomériques/pigments polymérisés et de la concentration en tanins précipitables par la méthode de dosage d'Adams-Harbertson**

Cette méthode de dosage se base sur les propriétés des anthocyanes à se décolorer en présence de métabisulfite, ainsi que sur les propriétés des pigments polymérisés et des tanins à précipiter avec les protéines (275). Ainsi, cette analyse sépare et quantifie les pigments monomériques, les petits pigments polymérisés, les grands pigments polymérisés et les tanins condensés précipitables avec l'albumine de sérum bovin. Il a été démontré que cette protéine se lie préférentiellement avec les proanthocyanidines dont le degré de polymérisation est supérieur à 3 (335). Les petits pigments polymérisés représentent donc les dimères et les trimères d'anthocyanes et flavan-3-ols, et les grands pigments polymérisés toutes les chaînes supérieures à 3 unités monomères. 125 µL de vin sont dilués dans une solution modèle de vin contenant 12 % d'éthanol v/v, 5 g/L d'acide tartrique, et ajusté à pH 3,3 avec de l'hydroxyde de sodium. Après plusieurs tests, le choix de cette dilution permet d'obtenir une absorbance inférieure à 1,2 durant le processus d'analyse.

Dans un premier cône Eppendorf de 1,5 mL on mélange 500 µL de vin dilué et 1 mL d'une solution tampon acide acétique-chlorure de sodium (acide acétique à 200 mM et chlorure de sodium à 170 mM, ajusté à pH 4,9 avec de l'hydroxyde de sodium). L'absorbance à 520 nm (10 mm de parcours optique avec l'eau comme blanc) de 1 mL de cette solution est alors mesurée (valeur A). Ensuite, 80 µL d'une solution à 0,36 M de métabisulfite de sodium sont ajoutés. Après 10 min, l'absorbance à 520 nm est de nouveau mesurée (Valeur B).

Dans un second cône Eppendorf de 1,5 mL on mélange 500 µL de vin dilué et 1 mL de la solution tampon acide acétique-chlorure de sodium contenant 1 g/L d'albumine de sérum bovin (BSA). Après 20 min, le cône Eppendorf est centrifugé pendant 5 min à 13500 g. 1 mL du surnageant est alors mélangé à 80 µL de la solution à 0,36 M de métabisulfite de sodium. Après 10 min, l'absorbance à 520 nm est mesurée (Valeur C).

Les absorbances dues aux pigments monomériques (MP), aux petits pigments polymériques (SPP) et aux grands pigments polymériques sont obtenues par les formules suivantes :

$$A_{(MP)} = \text{Valeur A} - \text{Valeur B}$$

$$A_{(SPP)} = \text{Valeur C}$$

$$A_{(LPP)} = \text{Valeur B} - \text{Valeur C}$$

L'absorbance due au total des pigments polymériques  $A_{(PP)}$  est aussi ajoutée comme la somme de  $A_{(SPP)}$  et  $A_{(LPP)}$ .

Dans le second Eppendorf, le culot est débarrassé du liquide restant et rincé avec 250 µL de la solution tampon acide acétique-chlorure de sodium. Il est ensuite centrifugé pendant 1 min à 13500 g et le surnageant est éliminé. Le culot est alors redissout dans 875 µL d'une solution contenant 5 % m/v de triéthylamine et 5 % m/v de dodécylsulfate de sodium ajusté à pH 9,4 avec de l'hydroxyde de sodium. L'absorbance témoin est mesurée à 510 nm (10 mm de parcours optique avec l'eau comme blanc), et 125 µL d'une solution de chlorure de fer III sont ajoutés (10 mM de chlorure de fer III et 10 mM d'acide chlorhydrique dans l'eau). Après 20 min à température ambiante, l'absorbance de la réaction est mesurée à 510 nm. L'absorbance relative aux tanins précipitables est calculée selon la formule suivante, où le coefficient de 0,875 permet de contrebalancer la dilution due à l'ajout de la solution de chlorure de fer III :

$$A_{(tanins)} = A_{510\text{témoin}} - A_{510\text{réaction}} / 0,875$$

Cette absorbance est alors reportée sur une courbe de calibration pour la catéchine qui est préalablement établie, et permet ainsi d'exprimer une valeur en mg/L d'équivalent catéchine. Les avantages de doser les tanins précipitables de façon indirecte après réaction au chlorure de fer III sont d'une part de les dissocier des pigments polymérisés, et d'autre part d'avoir une réponse d'absorbance qui s'affranchit de la nature des monomères de flavan-3-ols et du degré de polymérisation (336).

Une représentation schématique de la méthode de dosage est présentée sur la Figure 36, comme proposée par Adams et Harbertson en 2003.

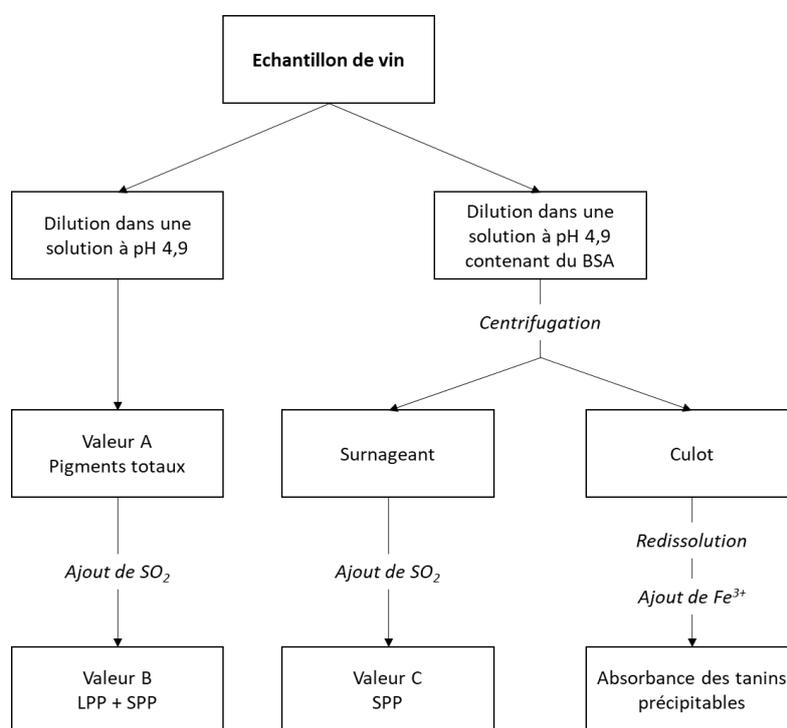


Figure 36 : Dosages des pigments et tanins précipitables (275)

## 5. Détermination de la concentration en tanins par précipitation à la méthylcellulose

Cette méthode de dosage se base sur les propriétés des tanins à précipiter avec des polysaccharides (327,337). Les tanins réagissent avec la méthylcellulose pour former un complexe tanin-polymère insoluble isolable par centrifugation. Dans un premier cône Eppendorf de 5 mL, 50 µL de vin sont mélangés à 800 µL d'une solution de sulfate d'ammonium saturée, 1950 µL d'eau et 1200 µL d'une solution de méthylcellulose (0,04 % dans l'eau, viscosité : 1500 cP, 2% dans l'eau à 20 °C). Dans un deuxième cône Eppendorf, 50 µL de vin sont mélangés à 800 µL d'une solution de

sulfate d'ammonium saturée et 3150 µL d'eau. Après homogénéisation et 10 min à température ambiante, les deux cônes Eppendorf sont centrifugés à 10000 g pendant 5 min. L'absorbance du surnageant est alors mesurée à 280 nm (10 mm de parcours optique avec l'eau comme blanc, dans une cuvette en quartz).

L'absorbance des tanins est définie comme la différence d'absorbance entre les deux valeurs obtenues, et celle-ci est reportée sur une courbe de calibration pour la catéchine. Le dosage des tanins est mesuré en mg/L d'équivalent catéchine.

Contrairement au dosage à l'albumine de sérum bovin précédemment présenté, cette valeur mesure aussi les pigments polymérisés qui pourraient précipiter, et apparaît donc comme un dosage global des tanins condensés comprenant aussi ceux liés aux anthocyanes.

## **6. Détermination des tanins totaux par hydrolyse acide (dosage de Bate-Smith)**

Cette méthode, développée par Bate-Smith, s'appuie sur les propriétés des proanthocyanidines à se transformer en anthocyanidines par chauffage en milieu acide (114). Dans deux tubes à hydrolyse fermés sont mélangés 4 mL de vin dilué 50 fois, 2 mL d'eau et 6 mL d'acide chlorhydrique à 37 %. Le premier tube est placé dans un bain-marie chauffé à 100 °C. Le second tube est placé dans un bain de glace à 0 °C. Après 30 min, 1 mL d'éthanol est ajouté aux deux tubes, et l'absorbance est mesurée à 550 nm (10 mm de parcours optique avec l'eau comme blanc, dans une cuvette en quartz). La différence d'absorbance entre les deux valeurs est alors reportée sur une courbe de calibration établie par Ribéreau-Gayon et Stonestreet, via la formule suivante :

$$[C_{\text{tanins totaux}}](\text{mg/L}) : 19330 \times (A_{550_{100^\circ\text{C}}} - A_{550_{0^\circ\text{C}}})$$

Cette méthode n'est utilisée que sur les vins finis commerciaux dénués de sucres, car le bain-marie dégrade les sucres et perturbe la lecture de la densité optique liés aux tanins totaux.

## **7. Détermination du degré de polymérisation moyen (DPM) par phloroglucinolyse**

### ***a) Principe général de la réaction de phloroglucinolyse***

Il est possible, par dépolymérisation des tanins condensés par phloroglucinolyse, de déterminer le degré de polymérisation moyen d'un échantillon. En milieu acide et sous l'action de la

chaleur, les liaisons interflavanes des proanthocyanidines peuvent être clivées. L'unité terminale se retrouve alors relarguée dans le milieu, alors qu'un adduit phloroglucinol nucléophile vient s'ajouter en position C4 de l'unité d'extension (Figure 37). Lorsque la réaction est totale, il est alors possible de calculer le degré de polymérisation moyen grâce au rapport entre les unités flavan-3-ols et les unités possédant un adduit phloroglucinol (109,277).

Une fois la réaction de dépolymérisation complète, les quatre flavan-3-ols majoritaires ainsi que les adduits phloroglucinol peuvent être identifiés :

- La catéchine (C) et la catéchine-phloroglucinol (C-P)
- L'épicatéchine (EC) et l'épicatéchine-phloroglucinol (EC-P)
- L'épigallocatechine (EGC) et l'épigallocatechine-phloroglucinol (EGC-P)
- L'épicatéchine-gallate (ECG) et l'épicatéchine-gallate-phloroglucinol (ECG-P)

La caractérisation de ces composés permet, en plus du D<sub>Pm</sub>, d'obtenir le ratio entre chaque unité, le pourcentage de galloylation et le pourcentage de prodelphinidine.

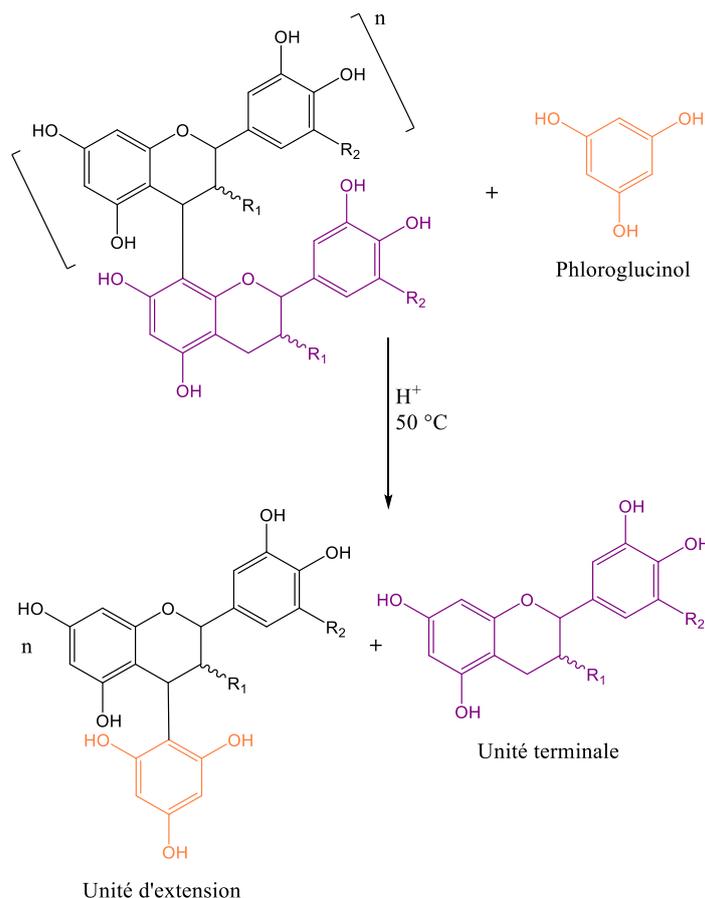


Figure 37 : Réaction de phloroglucinolyse sur les tanins condensés (277)

### ***b) Purification de l'échantillon***

Avant de pouvoir effectuer la réaction de phloroglucinolyse, une purification de l'échantillon est nécessaire pour éliminer tous les composés susceptibles d'interférer, comme les sucres, les acides ou l'éthanol. Pour cela, 5 mL d'échantillon sont évaporés à sec, puis repris dans 20 mL d'eau. La solution est alors déposée sur une cartouche  $C_{18}$  SPE (60 mL, 10 g), préalablement activée par un volume de méthanol (environ 50 mL) et un volume d'eau pour la préparer aux conditions initiales. L'échantillon est alors rincé par un volume d'eau pour éliminer les composés indésirables. La fraction de tanins condensés est ensuite éluée par un volume de méthanol et récupérée. Le volume obtenu est évaporé à sec et repris dans un volume de 2 mL de méthanol, avant d'être filtré (diamètre de pores de  $0,45\text{ }\mu\text{m}$ ) pour éliminer les résidus possibles de silice de la cartouche.

### c) Réaction de phloroglucinylation et analyse en CLHP-UV

100 µL de l'échantillon purifié précédemment sont placés dans un vial de 2 mL, auxquels on ajoute 100 µL de solution de phloroglucinol (50 g/L de phloroglucinol et 10 g/L d'acide ascorbique dilués dans du méthanol acidifié à 0,1 N d'acide chlorhydrique). Le milieu réactionnel fermé est alors placé dans une étuve à 50 °C pendant 20 min. A la fin de ce temps, 1 mL d'une solution d'acétate de sodium (40 mmol/L dans l'eau) est ajouté afin de stopper la réaction.

L'échantillon est alors injecté sur une CLHP Thermo-Scientific *Accela series*, avec comme longueur d'onde de détection 280 nm. La colonne ayant servi à la séparation est une phase inverse Xterra RP18 (100 × 4,6 mm, 3,5 µm) (Waters, USA) équipée d'une pré-colonne de même phase. Le débit d'injection est réglé à 1 mL/min, avec un volume d'injection de 20 µL à 15 °C. Les solvants utilisés sont (A) eau-acide acétique (99/1, v/v), et (B) méthanol, avec le gradient d'élution suivant (Tableau 10) :

Temps (min)	0	25	45	60	62	67	68	71
Solvant A (%)	95	95	80	68	0	0	95	95
Solvant B (%)	5	5	20	32	100	100	5	5

Tableau 10 : Gradient d'élution pour la séparation des flavan-3-ols et flavan-3-ol-phloroglucinol en CLHP

Les temps de rétention spécifiques à chaque composé recherché ont été préalablement identifiés par leur masse moléculaire d'ionisation spécifique (Tableau 11) :

Composé	Masse [M-H]
Catéchine	289
Catéchine-phloroglucinol	413
Epicatéchine	289
Epicatéchine-phloroglucinol	413
Epigallocatechine	305
Epigallocatechine-phloroglucinol	429
Epicatéchine-gallate	441
Epicatéchine-gallate-phloroglucinol	565

Tableau 11 : Masse moléculaire d'ionisation des flavan-3-ols et flavan-3-ol-phloroglucinol

Grâce au coefficient d'extinction molaire de chaque composé, leur concentration est obtenue en mmol/L (277). Ces teneurs permettent de calculer le degré de polymérisation moyen (DPM), le pourcentage de galloylation (% G) et le pourcentage de prodéphinidine (% P) via les formules suivantes :

$$DPm = \frac{\Sigma \text{ unités terminales} + \Sigma \text{ unités d'extensions}}{\Sigma \text{ unités terminales}}$$

$$\% G = \frac{EGC + EGC - P}{\Sigma \text{ unités terminales} + \Sigma \text{ unités d'extensions}}$$

$$\% P = \frac{EGC + EGC - P}{\Sigma \text{ unités terminales} + \Sigma \text{ unités d'extensions}}$$

## **D. Analyses sensorielles**

Afin de quantifier l'astringence et l'amertume d'un panel de vins étudiés (Tableau 8), une analyse sensorielle descriptive a été mise en place. L'évaluation de ces propriétés organoleptiques permet de rechercher les corrélations possibles entre la composition phénolique, l'analyse spectrale et l'appréciation sensorielle des vins.

Pour s'affranchir des biais de dégustation dû à l'environnement, l'ensemble des analyses sensorielles ont lieu dans la salle de dégustation de l'Institut des Sciences de la Vigne et du Vin. Il s'agit d'une salle blanche possédant un éclairage uniforme, une température contrôlée à 20 °C, et légèrement en surpression pour éviter la pénétration d'odeurs extérieures. Elle dispose également de postes de dégustation individuels équipés d'un évier crachoir et séparés par des plaques de plexiglass opaques pour éviter les interactions.

Tous les dégustateurs ayant participé aux analyses sensorielles sont habitués à la dégustation de vin dans le domaine de la recherche. Compte tenu de l'étalement des essais sur plusieurs mois, il n'a pas été possible de constituer un panel fixe. Cependant, chaque session s'est composée d'un nombre important de dégustateurs (de 18 à 28) avec une répartition la plus homogène possible entre hommes et femmes. L'évaluation de l'astringence et de l'amertume pouvant être parfois difficile à appréhender et sujette à la confusion, tous les dégustateurs ont suivi plusieurs séances d'entraînement.

### **1. Entraînement des dégustateurs**

Pour analyser l'astringence et l'amertume, il est important que ces deux notions soient bien comprises et différenciées par le dégustateur. Pour ce faire, le panel a suivi un premier entraînement avec une gamme de concentration croissante connue d'astringence et d'amertume. Les juges ont donc pu se familiariser avec ces perceptions et l'augmentation de leur intensité.

Pour modéliser l'astringence, une solution modèle de sulfate d'aluminium a été utilisée (338). 125, 250, 500, 1000 et 2000 mg de sulfate d'aluminium sont dissout dans 1 L d'eau à 12 % vol. d'éthanol.

Pour modéliser l'amertume, une solution de sulfate de quinine a été utilisée (338). 5, 10, 20, 30 et 40 mg de sulfate de quinine sont dissouts dans 1 L d'eau à 12 % vol. d'éthanol.

Ces deux types d'entraînements ont permis aux dégustateurs de travailler leur perception sur l'intensité de l'astringence et l'amertume. Les mêmes solutions modèles que précédemment ont aussi été utilisées, mais avec la non-connaissance des concentrations lors de sessions spécifiques. Les dégustateurs ont alors déterminé pour chaque gamme s'il s'agissait de l'astringence ou de l'amertume, et ont pu classer chaque verre du moins concentré au plus concentré sur une échelle non structurée. Ces deux dernières séances ont aussi permis de déterminer les concentrations considérées comme moyennes pour l'astringence et l'amertume par le panel pour chaque gamme, à savoir 1000 mg/L pour le sulfate d'aluminium et 20 mg/L pour le sulfate de quinine.

## **2. Dégustation des vins de l'étude**

La disposition de la dégustation des vins étudiés a été proposée pour éviter toutes influences extérieures sur la perception, ainsi que certains biais possibles dans ce genre d'étude :

- Le nombre d'échantillons présentés était inférieur ou égal à huit, pour éviter au maximum le phénomène d'addition notamment perçu pour l'astringence.
- Les vins étaient présentés en verre opaque de type verre INAO, pour éviter l'influence de la couleur sur la perception, et étaient identifiés par un code à 3 chiffres aléatoires.
- Un seul verre était proposé par poste de dégustation, pour deux raisons : la première, pour éviter la comparaison entre échantillons, et la seconde, pour obliger le dégustateur à marquer une pause entre chaque dégustation et éviter les effets d'additions.
- L'ordre de dégustation a été randomisé pour chaque juge.
- Chaque poste disposait d'un verre d'eau pour pouvoir se rincer la bouche, et d'un biscuit pour diminuer la réminiscence de l'astringence et permettre de se rapprocher de la neutralité buccale.

- A chaque début de séance de dégustation, les juges devaient tester une solution modèle d'astringence et d'amertume moyenne définie précédemment. Ces solutions ont pour effet de donner un repère stable et précis entre chaque séance.

La disposition d'un poste de dégustation type est présenté en Figure 38.



Figure 38 : Poste de dégustation type proposé pour définir l'astringence et l'amertume des vins étudiés

Pour caractériser l'astringence et l'amertume, il a été demandé aux dégustateurs de noter la perception en bouche sur une échelle non structurée allant de peu intense à très intense (Figure 39). Cette échelle mesurant 10 cm, nous avons ensuite mesuré la distance séparant le début de l'échelle de la notation du dégustateur, et attribuer au vin une valeur d'astringence et d'amertume comprise entre 0 et 10. Ils ont aussi pu renseigner la qualité tanique des vins, de fin à grossier, comme complément d'information à la perception d'astringence.

Nom : .....		Date : .....	
<b>Dégustation</b>			
Après avoir dégusté les solutions de référence, déguster les vins dans l'ordre proposé et évaluer l'intensité de l'astringence et l'amertume perçue, ainsi que la qualité tannique.			
<b>Vin 451</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier
<b>Vin 054</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier
<b>Vin 737</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier
<b>Vin 642</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier
<b>Vin 079</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier
<b>Vin 285</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier
<b>Vin 883</b>			
Astringence	Peu intense	_____	Très intense
Amertume	Peu intense	_____	Très intense
Qualité tannique	Fin	_____	Grossier

Figure 39 : Feuille de dégustation type proposée pour définir l'astringence et l'amertume des vins étudiés

## E. Analyses des données par régression PLS

Avant de construire les modèles de prédiction par régression PLS, une analyse préliminaire du jeu de données est nécessaire. Cela permet d'observer la variabilité, les groupes de corrélations possibles entre échantillons ou variables, ainsi que les possibles valeurs aberrantes qu'il faudrait écarter de la calibration. Des analyses en composantes principales ont donc été effectuées, en utilisant le logiciel libre R. Des fonctionnalités supplémentaires ont été ajoutées : l'environnement de développement RStudio ainsi que le package FactomineR.

### 1. Prétraitement des spectres

Pour éviter de dénaturer de façon excessive les spectres utilisés, un seul prétraitement a été appliqué : le centrage par la moyenne (*mean centering*). Le principal avantage d'avoir une seule et unique transformation est de pouvoir facilement comparer les résultats de prédiction entre eux en

s'affranchissant des biais induits par des algorithmes plus complexes. Cela permet aussi d'éviter un maximum les sur-corrélations, c'est-à-dire d'assimiler les données étudiées avec une réponse spectrale trop transformée qui ne correspond pas au composé recherché. Le centrage par la moyenne est un moyen efficace de mettre en évidence les zones de variabilité du jeu de données.

Sur la Figure 40 nous présentons les spectres UV-visible et IRTF du panel de vins finis commerciaux après un centrage par la moyenne. En comparant ces spectres avec ceux non centrés par la moyenne (Figure 33 et Figure 34), nous constatons une mise en évidence de la variabilité. Cette transformation permet de voir les zones présentant les plus fortes variations dans les spectres, qui n'étaient pas forcément visibles auparavant.

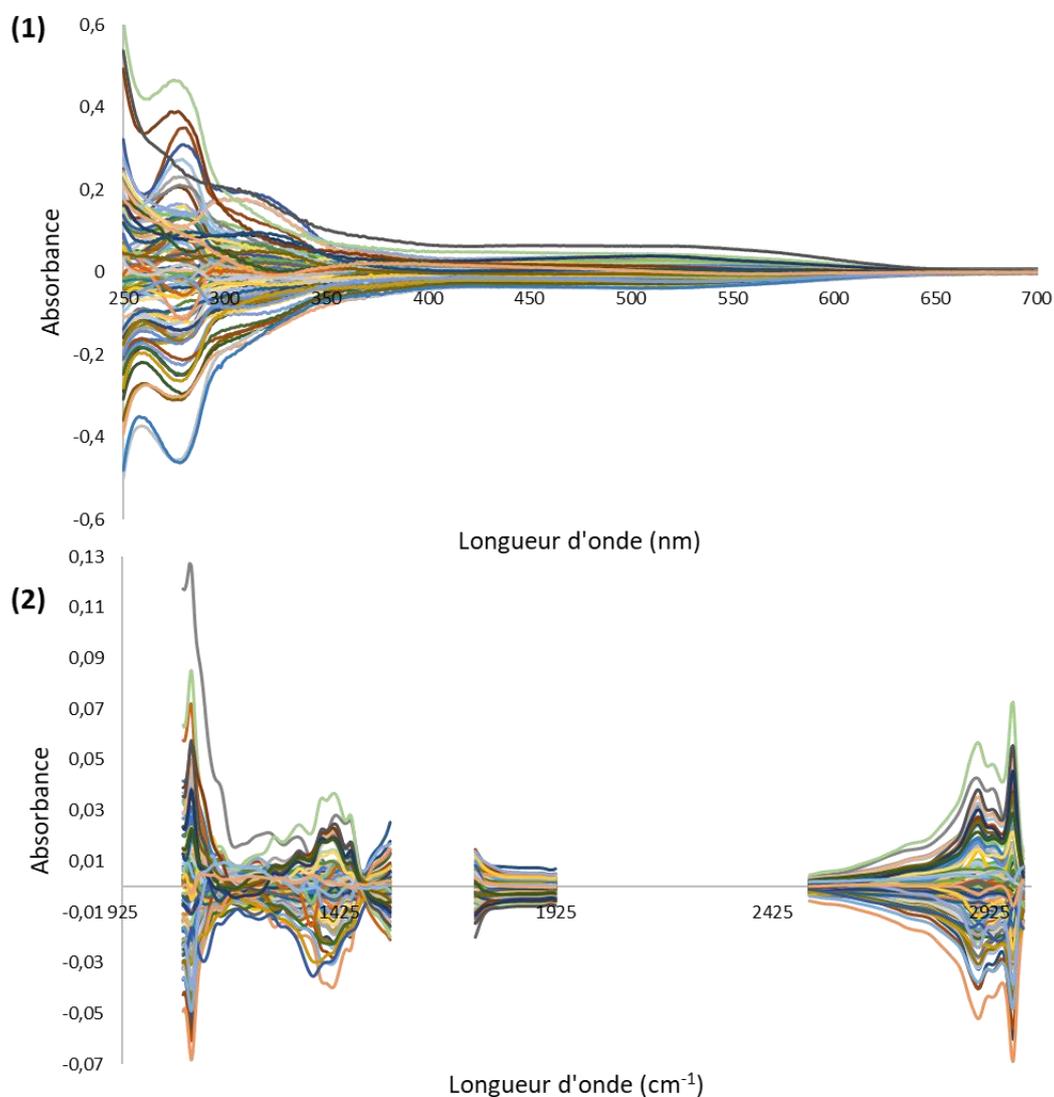


Figure 40 : (1) Spectres UV-visible et (2) spectres IRTF du panel de vins finis commerciaux après centrage par la moyenne

## 2. Conception de modèles de prédiction par régression PLS

### a) Logiciels et algorithmes

Afin de concevoir les modèles de prédiction cherchant à corrélérer les analyses de références aux spectres UV-visible et IRTF, l'environnement de développement Matlab 2017 (MathWorks, Natick, MA, USA) a été utilisé. Il a été couplé au logiciel d'extension PLS\_Toolbox (Eigenvector, Manson, WA, USA) qui permet d'apporter à Matlab les algorithmes de régression PLS ainsi qu'une interface graphique pour la conception des modèles.

L'algorithme utilisé est le SIMPLS. Il calcule les facteurs PLS directement sous forme de combinaisons linéaires des variables d'origine, et permet d'éviter certains biais observés sur l'algorithme (NIPALS)-PLS classiquement utilisé (339).

### b) Choix de la validation croisée et du jeu de données pour la validation

Le choix de la validation croisée est le *leave-p-out cross-validation*, où le jeu de données est divisé en un nombre  $p$  de sous-groupes qui vont servir à tester la prédiction du modèle. Le nombre de groupes par défaut utilisé est de dix. Il existe plusieurs méthodes pour définir ces sous-groupes, celle qui a été utilisée est la méthode dite *venetian blinds*. Le premier échantillon est réparti dans le groupe 1, le deuxième échantillon dans le groupe 2, etc. Arrivé à l'échantillon  $p+1$ , celui-ci est ajouté au groupe 1, et cette disposition continue jusqu'à la répartition complète du jeu de données. Cette méthode permet d'éviter la division par bloc sans interruption, et ainsi d'éviter de grouper ensemble des échantillons proches. Ainsi, chaque sous-groupe cherche à être représentatif du jeu de données complet.

Lorsque le jeu de données comporte un nombre assez important d'échantillons, il peut être divisé en un groupe de calibration et un groupe de validation. Pour cette répartition, l'algorithme de Kennard-Stone a été utilisé (340). Il permet de maximiser la cohérence entre les erreurs d'étalonnage et de validation, et donc d'obtenir un set de validation représentatif de la variabilité du jeu de données. L'avantage principal en comparaison à un partitionnement aléatoire est qu'il dispose d'une répétabilité accrue. Cela évite donc le biais correspondant au set de validation aléatoire qui est hautement prédictible et ne reflète pas la réelle robustesse du modèle.

### c) Sélection du nombre de variables latentes

Le choix du nombre de variables latentes est fait pour minimiser l'erreur de validation croisée et/ou l'erreur de prédiction. Cependant, lorsque la diminution de cette erreur est minime, il est préférable de se diriger vers un nombre de variables latentes plus faible et éviter une sur-corrélation. Afin de bien comprendre comment le choix des variables latentes a été effectué, un exemple est donné (Figure 41). L'erreur quadratique moyenne de calibration (RMSEC) et de validation croisée (RMSECV) est représentée en fonction du nombre de variables latentes.

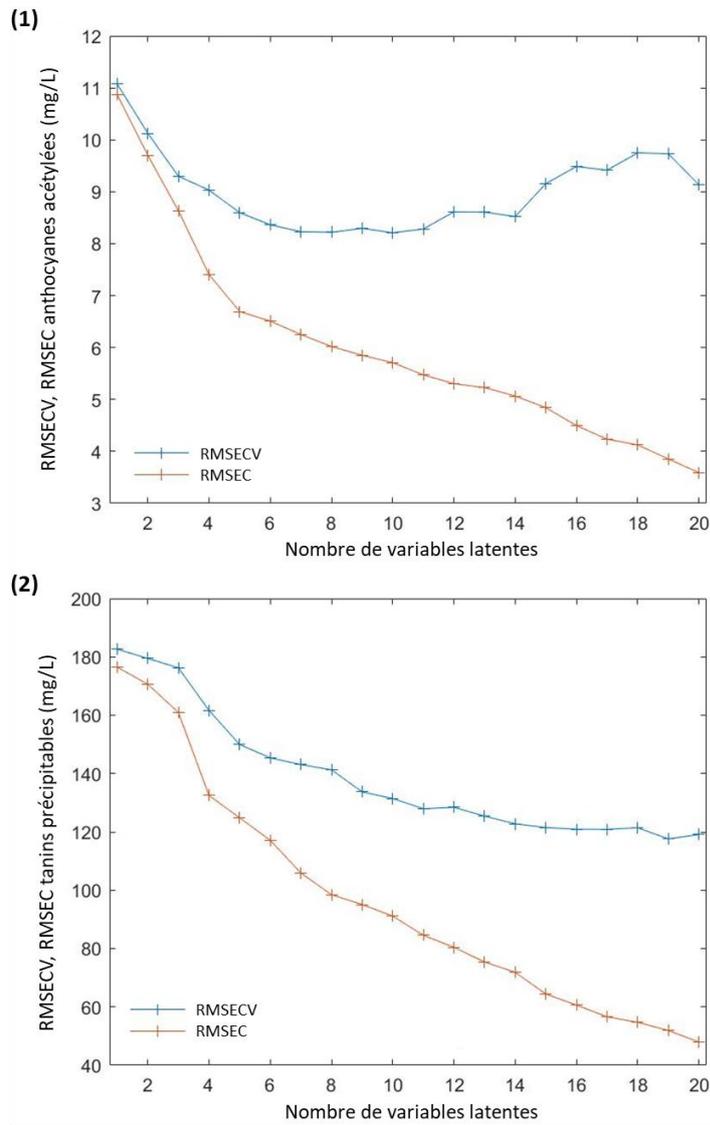


Figure 41 : Erreur quadratique moyenne de calibration (RMSEC) et de validation croisée (RMSECV) en fonction du nombre de variables latentes : exemple de modèles IRTF sur un panel de vins finis commerciaux.

Dans l'exemple (1), un modèle de prédiction IRTF des anthocyanes acétylées sur un panel de vins finis commerciaux, l'erreur de validation décroît fortement dans un premier temps, stagne, puis augmente. Le choix le plus judicieux en termes de variables latentes se situe après la forte diminution, soit 5 variables. Il est aussi possible de choisir 7 variables, mais le gain de robustesse apparaît minime.

Dans l'exemple (2), un modèle de prédiction IRTF des tanins précipitables sur un panel de vins finis commerciaux, la meilleure validation croisée est donnée pour 19 variables latentes. Cependant, nous constatons une légère augmentation de l'erreur après 11 variables, il sera donc déconseillé de choisir un nombre de variables latentes supérieur.

A partir du moment où une augmentation de l'erreur de validation est observable, le gain possible de précision grâce à un nombre plus élevé de variables latentes peut être dû à une surexploitation des données, et est donc à proscrire.

---

## **III. Résultats et discussions**

---



Les objectifs principaux de ce travail portent sur la prédiction et l'appréciation des tanins et des anthocyanes grâce à l'IRTF et l'UV-visible du raisin au vin. Pour y parvenir, il était nécessaire de couvrir la plus grande variabilité représentée, de la grappe jusqu'au vin prêt à être dégusté. Compte tenu des possibilités de diversité de la composition de la matrice étudiée, nous avons dans un premier temps essayé de construire des modèles de prédiction pour chaque étape de la conception d'un vin.

Les résultats présentés sont abordés en trois parties. Tout d'abord, la conception de modèles de prédiction est présentée pendant la maturité du raisin jusqu'à la vendange. Ensuite, nous avons étudié des modèles lors de la vinification, et plus particulièrement l'étape de la fermentation alcoolique et de macération car il s'agit d'un moment crucial dans l'extraction des anthocyanes et des tanins. Enfin, nous avons travaillé sur des modèles de prédiction des polyphénols pour les vins commerciaux.

Pour chacune de ces étapes, des modèles spécifiques de prédiction ont été créés pour chaque paramètre étudié à l'aide des spectres IRTF et UV-visible enregistrés. Ils ont été soumis à des tests de validation afin de mettre en évidence leur robustesse. Par ailleurs, de nouveaux modèles ont aussi été créés en combinant l'UV-visible et l'IRTF pour rechercher la possibilité d'obtenir des modèles plus performants par combinaison de ces deux méthodes spectroscopiques.

De plus, nous avons travaillé sur l'élaboration de modèles essayant d'englober l'ensemble de tous les échantillons, pour apprécier la possibilité de s'affranchir ou pas de la variabilité des matrices (raisin, moût, vin) pour la prédiction des anthocyanes et des tanins.

## **A. Prédiction des concentrations en anthocyanes et tanins du raisin pendant la maturité**

La maturité du raisin a été suivie à raison d'un échantillon de raisin par semaine pendant quatre à cinq semaines précédant les vendanges. Cette période est utilisée par l'Union de Producteurs de Saint-Emilion pour suivre la maturité technologique et phénolique des parcelles qu'ils vinifient, et reflète donc les besoins analytiques du milieu viticole.

Afin d'étudier la prédiction des concentrations en anthocyanes et en tanins présents dans le raisin pendant sa maturité, nous avons dans un premier temps étudié la variabilité du jeu de données ainsi que l'évolution dans le temps de chaque paramètre étudié. Il était primordial de s'assurer de la bonne répartition des valeurs obtenues pour s'assurer que les modèles créés ultérieurement sont représentatifs.

## 1. Analyse du jeu de données.

### a) *Evaluation de la variabilité des analyses de références*

Le jeu de données est représenté par 123 échantillons, répartis sur trois millésimes. Pour le premier millésime composé de 24 échantillons, seules les analyses Glories ont été effectuées. Sur les deux autres millésimes, les analyses Glories, les analyses des anthocyanes moléculaires, le suivi de la polymérisation des pigments (dosage d'Adams-Harbertson), ainsi que l'analyse des tanins par précipitation à la méthylcellulose et à l'albumine de sérum bovin ont été effectués. Dans le Tableau 12 la variabilité de ces analyses est représentée. Les données produites intègrent les millésimes 2017, 2018 et 2019 pour les analyses Glories, et les millésimes 2018 et 2019 pour le reste des paramètres. Aucun échantillon n'a été écarté. Plusieurs paramètres sont ainsi répertoriés :

- La moyenne des résultats obtenus, ainsi que le minimum (Min) et le maximum (Max).
- L'écart-type des données obtenues.
- Le coefficient de variation (CV) ou écart-type relatif qui est défini comme le rapport entre l'écart type et la moyenne, et exprimé en pourcentage. L'écart-type seul ne permet pas forcément d'évaluer la dispersion des données autour de la moyenne, cette donnée nous permet donc de mieux appréhender la dispersion générale du paramètre étudié.
- L'écart-type de l'analyse de référence (Ecart-type Ref) qui est calculé comme l'écart-type du triplicat d'analyse pour chaque échantillon. Cette donnée, comparée à l'écart-type, permet de juger de la pertinence de l'analyse et de relativiser les erreurs des modèles de prédiction qui peuvent être induites en partie par l'erreur standard de l'analyse de référence. De plus, si cette donnée s'avère trop proche de la valeur de l'écart-type de l'analyse, il est possible de remettre en question l'analyse de référence qui peut s'avérer non appropriée pour le suivi du paramètre étudié sur le jeu de données.

Dans le Tableau 12, le paramètre Total glucoside représente la somme des anthocyanes-3-*O*-glucoside, Total acétyl représente la somme des anthocyanes-3-*O*-acétylglucoside, Total coumaroyl la somme des anthocyanes-3-*O*-coumaroylglucoside et Total anthocyanes la somme de toutes les anthocyanes moléculaires de l'échantillon. Pour les tanins, le paramètre BSA représente les tanins précipitables à l'albumine de sérum bovin, et MCP les tanins précipitables à la méthylcellulose.

Paramètre	Moyenne	Min	Max	Ecart-type	CV	Ecart-type Ref
<b>Analyse Glories</b>						
ApH1 (mg/L)	1327,64	508,62	2361,07	384,15	29 %	29,01
ApH3,2 (mg/L)	604,00	284,03	999,69	139,46	23 %	14,10
RPT (Absorbance)	48,83	9,35	107,57	15,61	32 %	3,97
<b>Anthocyanes moléculaires</b>						
Dp-3-O-glc (mg/L)	14,77	0,26	34,83	7,06	48 %	1,25
Cy-3-O-glc (mg/L)	11,96	2,18	29,65	6,14	51 %	0,43
Pt-3-O-glc (mg/L)	14,88	0,46	32,33	6,08	41 %	0,89
Pn-3-O-glc (mg/L)	25,01	6,92	52,92	10,31	41 %	0,72
Mlv-3-O-glc (mg/L)	83,33	20,16	172,29	30,41	36 %	3,10
Dp-3-O-acglc (mg/L)	2,56	0,10	5,40	1,11	43 %	0,24
Cy-3-O-acglc (mg/L)	1,68	0,38	3,55	0,74	44 %	0,10
Pt-3-O-acglc (mg/L)	3,29	0,17	6,44	1,23	37 %	0,23
Pn-3-O-acglc (mg/L)	4,30	1,06	8,27	1,52	35 %	0,24
Mlv-3-O-acglc (mg/L)	23,39	3,77	47,42	9,78	42 %	0,98
Pn-3-O-cmglc (mg/L)	1,35	0,34	3,15	0,56	42 %	0,14
Mlv-3-O-cmglc (mg/L)	5,87	1,00	14,07	2,69	46 %	0,39
Total glucoside (mg/L)	149,95	41,00	287,09	51,57	34 %	6,30
Total acétyl (mg/L)	35,22	9,25	65,58	12,30	35 %	1,63
Total coumaroyl (mg/L)	7,21	1,83	16,57	3,06	42 %	0,48
Total anthocyanes (mg/L)	192,38	52,74	369,16	65,09	34 %	8,15
<b>Etat de polymérisation des pigments</b>						
MP (Absorbance)	0,88	0,30	1,56	0,25	28 %	0,09
SPP (Absorbance)	0,53	0,00	1,57	0,37	70 %	0,07
LPP (Absorbance)	0,19	0,00	1,09	0,27	139 %	0,08
PP (Absorbance)	0,73	0,07	1,71	0,38	52 %	0,09
<b>Tanins</b>						
BSA (mg/L)	254,52	0,00	803,29	188,40	74 %	18,16
MCP (mg/L)	1117,37	0,00	2397,31	539,59	48 %	105,38

Tableau 12 : Paramètres des analyses de références obtenus sur les échantillons de raisins pendant la maturité (RPT : richesse polyphénolique totale, Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, CV : coefficient de variation)

Considérant les valeurs obtenues pour les analyses Glories, nous constatons que celles-ci sont pour les données d'ApH1 proches du double des données d'ApH3,2. La déviation standard des analyses (Ecart-type Ref) est faible en comparaison à l'écart-type des données, représentant 8 % de l'écart-type pour ApH1 et 10 % pour ApH3,2. Avec des coefficients de variation respectivement de 29 % et 23 %, ces deux paramètres montrent une variabilité correcte et une erreur acceptable, et semblent adaptés pour la conception de modèles de prédiction. Concernant la richesse polyphénolique totale

(RPT), si la variabilité s'apparente aux deux paramètres précédents, nous constatons néanmoins une erreur standard de l'analyse plus importante, correspondant à près de 25 % de l'écart-type. Ainsi, cette valeur sera à prendre en compte lors de la conception du modèle PLS correspondant, car elle influencera la robustesse de la prédiction.

Si l'on s'intéresse aux données des analyses chimiques des anthocyanes moléculaires, nous constatons que pour une majorité d'entre elles, les concentrations moyennes sont faibles, inférieures à 15 mg/L. Seules la malvidine-3-*O*-glucoside, la péonidine-3-*O*-glucoside et malvidine-3-*O*-acétylglucoside indiquent des teneurs supérieures. Dans la littérature, des études similaires sur le raisin ou le vin indiquent des résultats proches, sans que les limites de détection en UV-visible ou IRTF ne soient discutées (12,329). Cela peut être attribué à la grande variabilité de la matrice et la superposition de la réponse spectrale induite par la variabilité des molécules, qui influencera fortement cette limite de détection (279,341). Lors de la conception de modèles de prédiction, ces faibles teneurs en anthocyanes seront donc à prendre en compte pour éviter des sur-corrélations. Nous constatons également que la malvidine-3-*O*-glucoside représente une majeure partie des anthocyanes en présence, avec une moyenne de 83,33 mg/L pour une moyenne d'anthocyanes totales à 192,38 mg/L. Le coefficient de variabilité des concentrations des différentes anthocyanes varie entre un minimum de 35 % pour la péonidine-3-*O*-acétylglucoside et un maximum de 51 % pour la cyanidine-3-*O*-glucoside. Ainsi, nous pouvons estimer une évolution de ces anthocyanes sensiblement différentes. Cette variabilité s'atténue lorsque nous regardons les résultats des anthocyanes totales, avec un coefficient de variabilité de 34 %.

Parmi les paramètres qui permettent de mesurer le ratio entre pigments monomériques et polymériques on trouve de plus grandes disparités. Les pigments monomériques montrent une variabilité similaire à celle d'ApH1 et ApH3,2 (CV de 28 %), ce qui peut s'expliquer par une analyse par décoloration au bisulfite très similaire. Cependant, tout l'intérêt de la mesure des pigments monomériques réside dans sa comparaison avec les teneurs en petits et grands pigments polymérisés (SPP et LPP). Pour les échantillons étudiés, la variabilité de ces pigments polymérisés semble très forte, avec un coefficient de variation de 70 % pour les SPP et de 139 % pour les LPP, et semble indiquer une évolution du ratio pigment monomérique/polymérique importante. Il faut cependant prendre aussi en compte un écart-type de l'analyse de référence élevé en comparaison à l'écart-type du jeu de données.

Les deux paramètres mesurant les tanins précipitables sont intéressants à comparer, car ils montrent une variabilité très différente. Les tanins précipitables à l'albumine de sérum bovin (BSA) présentent un coefficient de variation de 74 %, et pourrait donc indiquer une évolution forte de la réactivité des tanins pendant la maturité. A l'inverse, les tanins précipitables à la méthylcellulose (MCP), avec une sélectivité du degré de polymérisation moyen moins importante que les protéines, présentent une variabilité moins importante (CV de 48 %).

Sur le jeu de données des échantillons étudiés, la variabilité de tous les paramètres apparaît assez importante pour indiquer une évolution de la teneur en anthocyanes et en tanins. Cependant, ces résultats seuls ne permettent pas d'affirmer une répartition homogène des données pour obtenir des modèles de prédiction représentatifs de l'évolution de la maturité. Dans ce but, il est nécessaire d'utiliser l'analyse en composantes principales afin de mettre en évidence les échantillons déviants ou la formation de groupe d'échantillons qui pourrait influencer les résultats de prédiction.

### ***b) Analyse en composantes principales (ACP) des analyses de références et des données spectrales***

#### ***(1) ACP des variables des analyses de références***

L'analyse en composantes principales (ACP) est un outil efficace pour évaluer la distribution globale d'un jeu de données où les individus sont représentés par plusieurs variables. Elle permet aussi de mettre en évidence des corrélations entre individus et entre variables. Pour concevoir cette analyse, il est nécessaire que tous les échantillons partagent les mêmes paramètres étudiés. Ainsi, les ACP ont été construites sur les millésimes 2018 et 2019. Nous nous sommes intéressés dans un premier temps à l'ACP des variables des analyses de références (Figure 42). Les données obtenues pour chaque monomère d'anthocyane ne sont pas utilisées pour la conception de cette ACP, afin d'éviter une surreprésentation de paramètres inhérents à ces composés qui pourrait influencer trop fortement l'ACP.

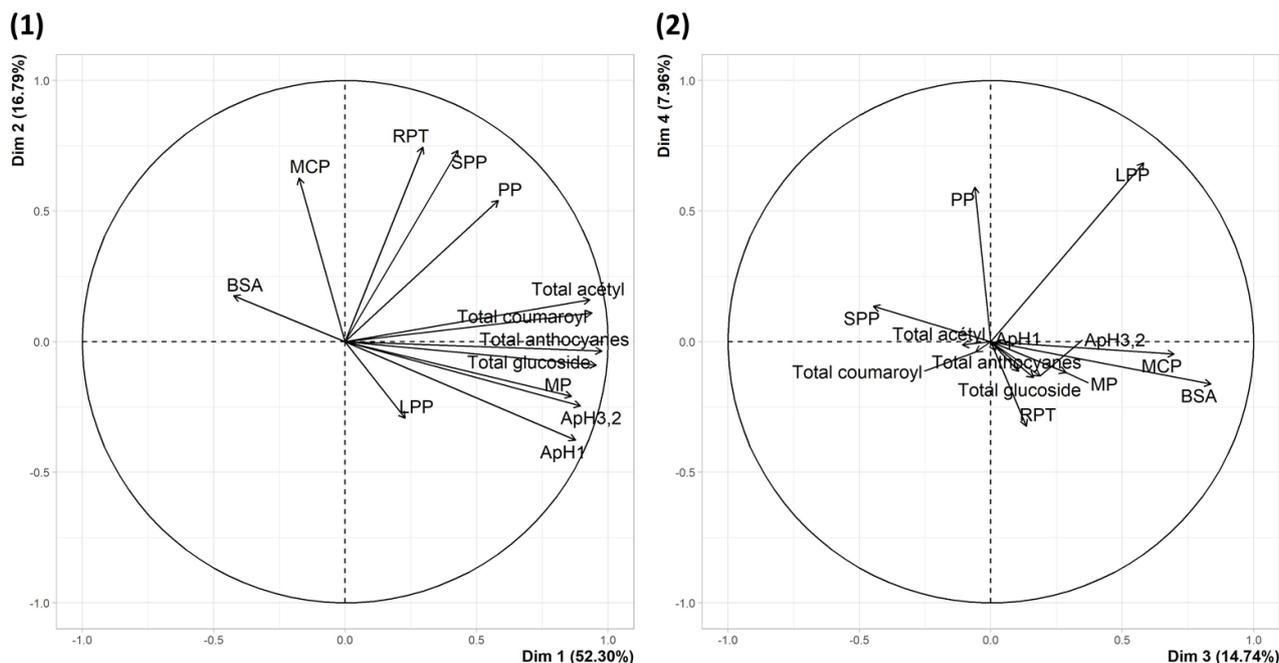


Figure 42 : ACP des variables des analyses de références : (1) dimensions 1 et 2, (2) dimensions 3 et 4 (RPT : richesse polyphénolique totale, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose)

Les deux premières dimensions de cette ACP expliquent un total de 69,09 % de la variance des analyses de références, avec une variance très forte sur le premier axe (52,30 %). Ces deux dimensions représentent donc fortement le jeu de données d'analyses étudié. Chaque paramètre est représenté par une flèche. Lorsque cette flèche est proche du cercle des corrélations, cela signifie que la variabilité représentée sur les deux dimensions explique la variabilité du paramètre. En observant cette ACP, nous constatons un groupe de variables fortement corrélé avec la dimension 1, composé d'ApH1, ApH3,2, les anthocyanes-3-O-glucoside (Total glucoside), les anthocyanes-3-O-acétylglucoside (Total acétyl), les anthocyanes-3-O-coumaroylglucoside (Total coumaroyl) et les anthocyanes totales. Lors de la maturité du raisin, la variabilité en polyphénols la plus importante est donc expliquée par la composition anthocyanique.

Deux paramètres apparaissent corrélés avec la dimension 2, la richesse polyphénolique totale (RPT) et les tanins précipitables à la méthylcellulose (MCP). La variabilité de cette dimension peut donc être attribuée aux concentrations en tanins condensés, mesurées par MCP et qui contribuent majoritairement aux valeurs de RPT. On note cependant que l'analyse des tanins précipitables à l'albumine de sérum bovin (BSA) montre une corrélation faible avec cette dimension. Si l'on regarde les résultats obtenus pour les dimensions 3 et 4, nous constatons une forte corrélation de BSA avec la

dimension 3, qui explique 14,74 % de la variance des données. Ainsi, la réactivité des tanins condensés aux protéines possède une évolution lors de la maturité qui lui est propre, qui semble peu corrélée avec celle des tanins précipitables à la méthylcellulose.

Afin de s'intéresser plus particulièrement aux anthocyanes moléculaires, nous avons analysé par ACP la variabilité des différentes anthocyanes (Figure 43).

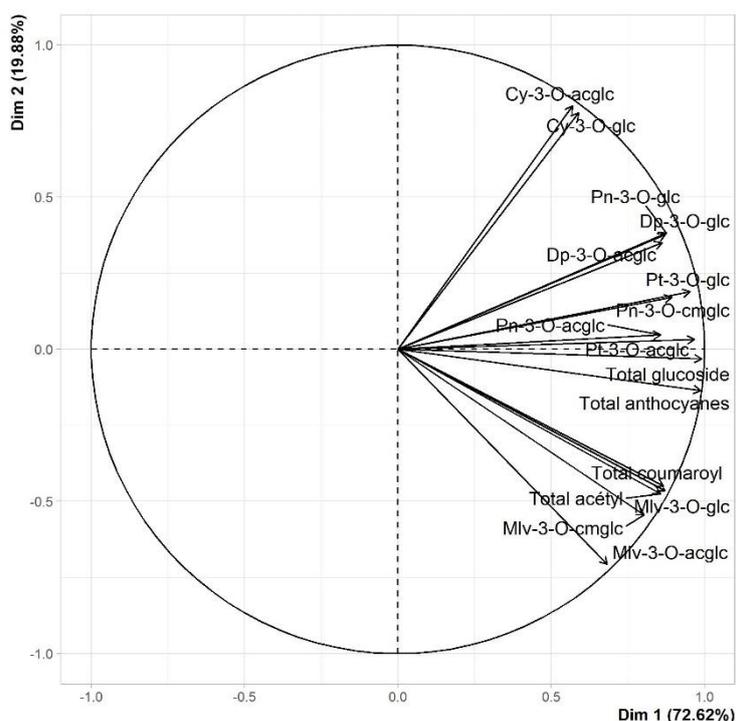


Figure 43 : ACP des variables des anthocyanes moléculaires, dimension 1 et 2 (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside)

L'on constate que la plus grande partie de la variance est expliquée par la dimension 1 (72,62 %). L'évolution de chacune des anthocyanes semble donc corrélée avec l'évolution des anthocyanes totales, avec toutefois quelques différences notables si l'on s'intéresse à la dimension 2. Il existe un regroupement par anthocyanes composées de la même anthocyanine, avec un groupe malvidine, un groupe cyanidine et un dernier groupe composé par la delphinidine, péonidine et pétunidine. Si l'on observe une distinction franche entre les anthocyanes composées de malvidine et cyanidine, nous constatons toutefois que le reste des anthocyanes est fortement corrélé avec la concentration en anthocyanes totales. De plus, la dimension 2 n'explique que 19,88 % de la variance, et tous les paramètres montrent une corrélation importante avec la dimension 1. Ces valeurs seront à prendre en

compte lors de la discussion des modèles, afin d'éviter d'interpréter des résultats qui pourraient être dus à la modélisation des anthocyanes totales et non à une anthocyane moléculaire.

Globalement ces analyses des variables permettent de mettre en évidence une variabilité idéale pour la conception des modèles, avec trois groupes de variables distinctes : les concentrations en anthocyanes, les concentrations en tanins condensés et les concentrations en tanins réactifs aux protéines. Il est maintenant nécessaire d'analyser par ACP les individus (échantillons) en fonction des analyses de références et des spectres, pour s'assurer d'une bonne homogénéité de leur disposition. Cela doit permettre également de mettre en évidence les échantillons déviants ainsi que des regroupements possibles entre échantillons.

*(2) ACP des individus en fonction des analyses de références et des spectres UV-visible et IRTF*

L'ACP des individus en fonction des analyses de références est représentée en Figure 44. Les données utilisées sont identiques à celles utilisées pour la Figure 42, les ACP des variables et des individus sont donc comparables entre elles. Chaque point représente un échantillon, et les données utilisées sont la moyenne du triplicat d'analyse.

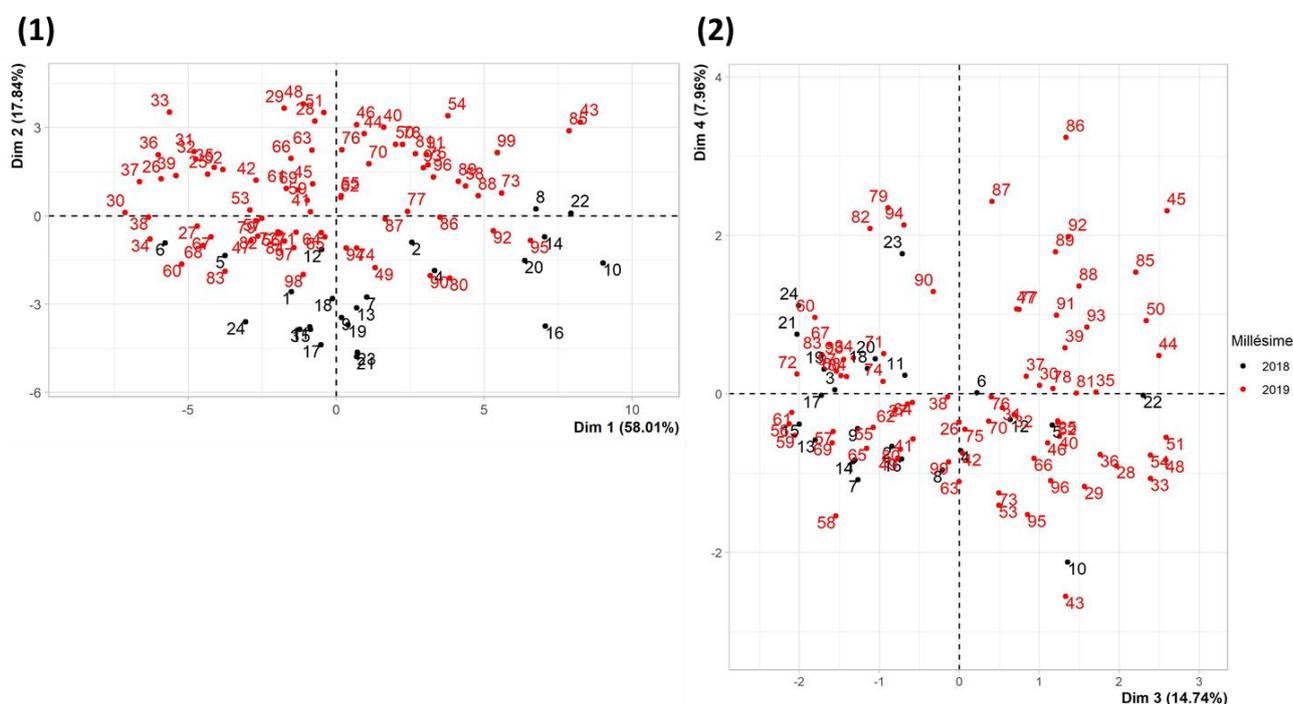


Figure 44 : ACP des individus des analyses de références : (1) dimensions 1 et 2, (2) dimensions 3 et 4

La première information notable sur l'ACP de la Figure 44 est un regroupement par millésime. Nous constatons une séparation par rapport à la dimension 2, corrélée aux valeurs de MCP et RPT. Le millésime impacte donc ces paramètres. Pour la conception de modèles de prédiction, ce type de discrimination n'est pas recherché car cela peut signifier une variabilité propre à chaque millésime et donc ne pas permettre aux modèles d'être prédictifs d'une année à l'autre. Cependant, nous remarquons que cette différence n'est notable que sur la dimension 2, ce qui signifie que les paramètres fortement corrélés aux autres dimensions ne présentent pas cette discrimination. Ainsi, le millésime ne semble pas impacter les paramètres relatifs aux anthocyanes ainsi que les tanins précipitables à l'albumine de sérum bovin.

En s'intéressant à la disposition du nuage de points des individus, nous ne constatons pas de regroupement notable où d'échantillons à l'écart. Il s'agit du type de dispersion recherché, car cela signifie que les échantillons représentent bien la variabilité des paramètres étudiés. De plus, l'absence d'échantillons déviants signifie que tout le jeu de données pourra être utilisé pour la conception des modèles, sans écarter de résultats.

L'analyse par ACP a aussi été utilisée sur les spectres UV-visible et IRTF, afin de s'assurer également d'une bonne homogénéité des échantillons, et de rechercher la possible distinction entre millésime. L'ACP des individus des spectres UV-visible est représentée en Figure 45.

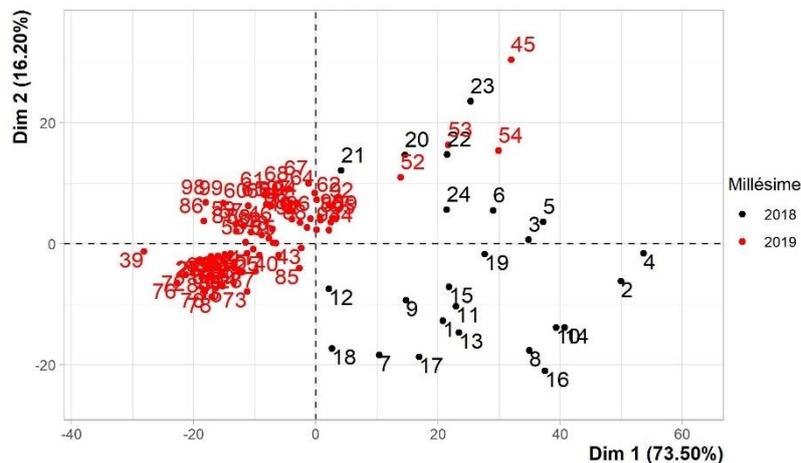


Figure 45 : ACP des individus des spectres UV-visible

On constate sur l'ACP de la Figure 45 la même séparation par millésime, sauf que contrairement à l'ACP des analyses de références, elle apparaît bien plus marquée et corrélée à la dimension 1, qui représente 73,50 % de la variance totale des spectres UV-visible. Le millésime impacte donc fortement l'empreinte spectrale, et plusieurs millésimes sont donc nécessaires pour concevoir un

modèle de prédiction robuste. De plus, nous remarquons un fort regroupement concernant les données du millésime 2019, qui montre une variabilité limitée en comparaison du millésime 2018.

Si l'on s'intéresse à l'ACP des spectres IRTF, nous observons également un regroupement du millésime 2018 dans le nuage de points (Figure 46). Cette distinction est cependant moins importante que sur l'ACP des spectres UV-visible, et la répartition des échantillons est bien plus homogène.

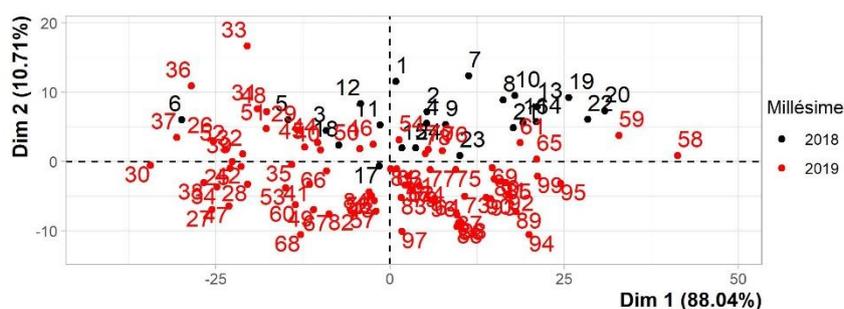


Figure 46 : ACP des individus des spectres IRTF

Ces différentes ACP des individus nous ont permis de mettre en évidence un fort effet millésime, qui semble impacter les résultats des tanins précipitables à la méthylcellulose et la richesse polyphénolique totale. Cette différence est aussi notable sur les spectres UV-visible et IRTF, et particulièrement marquée sur les spectres UV-visible. Afin de déterminer si cet effet millésime peut être impactant sur la conception de modèles de prédiction PLS, nous avons utilisé l'analyse discriminante PLS pour voir s'il était possible de différencier les millésimes en fonction des analyses de références, des spectres UV-visible et des spectres IRTF.

### ***c) Analyse discriminante PLS pour la différenciation des millésimes étudiés pendant la maturité du raisin***

Contrairement à la régression PLS, qui cherche à prédire un paramètre quantitatif en fonction d'un jeu de données, l'analyse discriminante PLS (PLSDA) permet de modéliser la distribution des individus en fonction de groupes définis au préalable. Nous avons donc utilisé cette méthode pour observer s'il était possible de différencier les échantillons en fonction de leur millésime. En Figure 47 est représentée l'analyse PLSDA en fonction des analyses de références. Sur cette représentation graphique l'échelle de prédiction est arbitraire et permet d'attribuer une valeur à chaque échantillon. Celle-ci permet ensuite de distribuer chaque individu en fonction de la limite de discrimination. Cette limite est déterminée pour maximiser la différenciation entre les groupes.

Pour toutes les analyses PLSDA présentées, les valeurs utilisées pour chaque échantillon sont la moyenne du triplicat d'analyse. Le prétraitement des données est un centrage par la moyenne. Les paramètres de validation croisée sont les suivants : 10 groupes sont définis par une répartition de type *venetian blinds*.

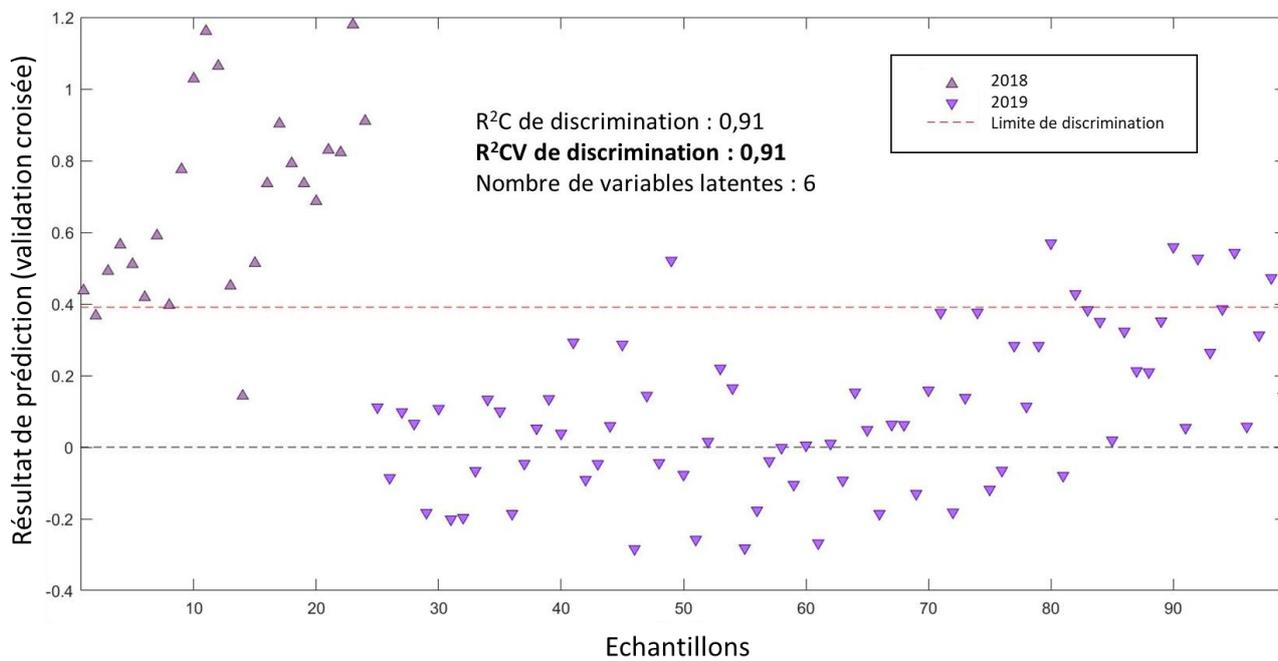


Figure 47 : Analyse PLSDA des millésimes en fonction des analyses de références ( $R^2C$  : coefficient de régression de calibration,  $R^2CV$  : coefficient de régression de validation croisée)

Nous constatons sur cette représentation graphique une discrimination par millésime très marquée, mais qui montre cependant des incertitudes. Ainsi, si une distinction apparaît, il n'est pas possible de prédire correctement une partie des échantillons en fonction des analyses de références. Néanmoins, cette discrimination montre un coefficient de régression de validation croisée élevé ( $R^2CV$  : 0,91), et montre donc des variations de la concentration en anthocyanes et en tanins fortement impactées par le millésime.

Nous avons remarqué précédemment que cette discrimination était aussi notable sur les spectres UV-visible et IRTF. Il est donc nécessaire d'analyser par PLSDA ces spectres (Figure 48).

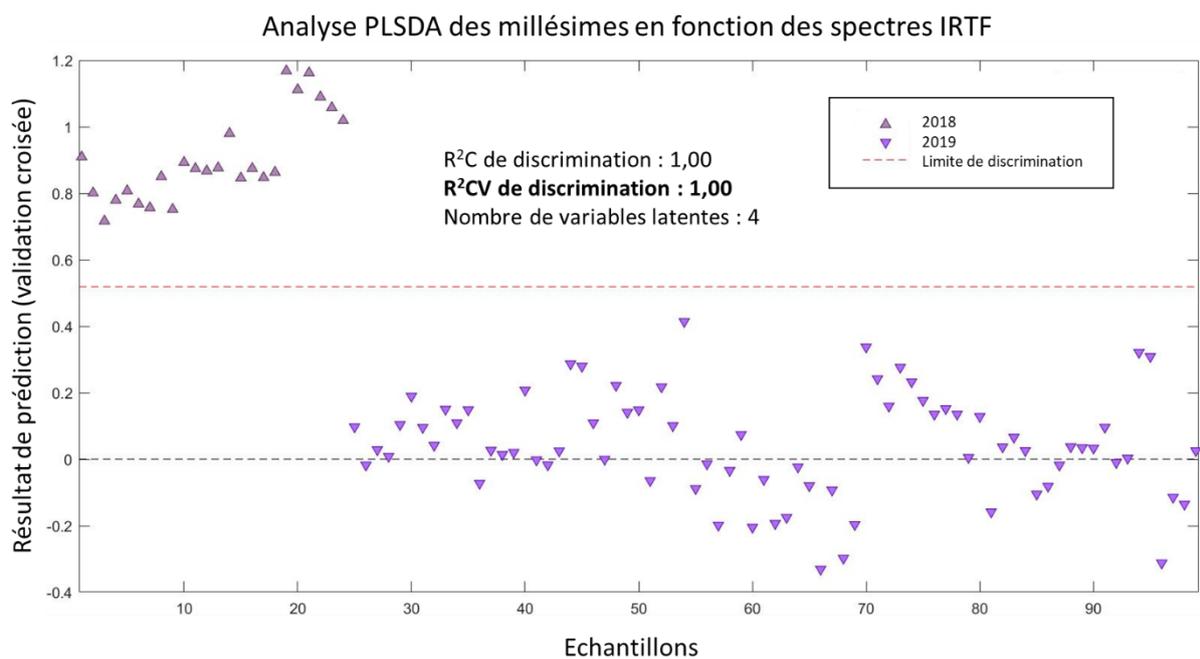
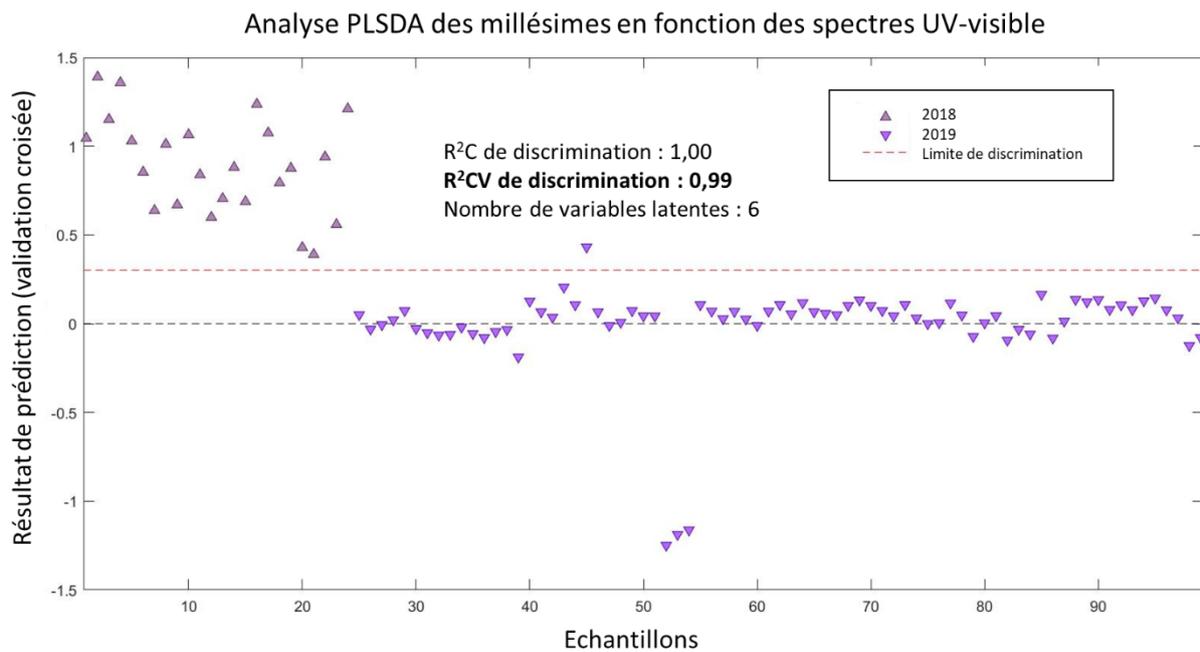


Figure 48 : Analyse PLSDA des millésimes en fonction des spectres UV-visible et IRTF (R<sup>2</sup>C : coefficient de régression de calibration, R<sup>2</sup>CV : coefficient de régression de validation croisée)

Premièrement, l'analyse PLSDA des millésimes en fonction des spectres UV-visible montre une discrimination forte, avec un seul échantillon qui n'est pas distribué dans le bon millésime (R<sup>2</sup>CV : 0,99). En comparant ces résultats à ceux de l'ACP des spectres UV-visible, nous observons le même style de répartition : le millésime 2018 montre une répartition hétérogène, alors que les échantillons

du millésime 2019 montrent des valeurs proches, à l'exception de 4 échantillons (45, 52, 53 et 54). Si l'on s'intéresse à l'ACP des analyses de références, nous n'observons pas la même démarcation pour ces échantillons, qui n'apparaît donc pas induite par les teneurs en anthocyanes ou en tanins. Deux hypothèses peuvent être formulées pour expliquer cette distinction : un effet matrice sur ces échantillons qui influence fortement l'empreinte spectrale, ou une erreur analytique lors de l'enregistrement du spectre. Dans ce dernier cas, les échantillons apparaîtront déviants lors de la conception de modèles de prédiction et pourront être écartés du jeu de données.

Sur l'analyse PLS-DA en fonction des spectres IRTF, les résultats de validation croisée indiquent une discrimination parfaite ( $R^2_{CV} : 1,00$ ). Cette modélisation permet donc de séparer sans erreurs les deux millésimes étudiés. Dans le but de concevoir des modèles de prédiction utilisables d'un millésime à l'autre, ce type de répartition n'est pas recherchée. En effet, pouvoir discriminer chaque millésime signifie que ceux-ci possèdent une empreinte spectrale spécifique, qui peut influencer sur la lecture des paramètres recherchés. A contrario, s'il est possible de prédire ces paramètres en utilisant un millésime pour la calibration et l'autre pour la validation, cela signifie que le modèle est robuste et s'affranchit de la variabilité de la matrice.

La différenciation entre millésimes est observable sur les analyses de références ainsi que sur les spectres, et peut donc être induite par les teneurs en anthocyanes et en tanins qui ont été analysées. Cependant, les effets des autres composés du raisin sur la matrice ne sont pas à négliger. Afin de s'assurer de la reproductivité des prédictions des modèles PLS construits, l'étape de validation sera primordiale. Il sera donc nécessaire de valider les modèles PLS d'un millésime à l'autre, pour s'assurer de leur robustesse dans la prédiction de données extérieures à la calibration.

## **2. Construction de modèles PLS pour la prédiction des anthocyanes et des tanins pendant la maturité du raisin**

Nous avons pu suivre l'évolution des tanins et des anthocyanes du raisin pendant la maturité jusqu'aux vendanges grâce à différentes analyses de références. Avant de discuter des résultats des modèles de prédiction, il est important de s'intéresser à l'évolution des paramètres étudiés lors de la maturité, afin de juger de l'évolution globale des tanins et des anthocyanes du raisin ainsi que de la pertinence de chaque analyse.

***a) Construction de modèles de prédiction PLS pour les analyses Glories sur les millésimes 2017, 2018 et 2019 pendant la maturité du raisin***

***(1) Evolution des paramètres Glories pendant la maturité du raisin***

La méthode Glories est un dosage qui permet d'obtenir le potentiel en anthocyanes extractibles (ApH3,2), le potentiel total en anthocyanes (ApH1), ainsi que la richesse polyphénolique totale (RPT). À partir de ces trois valeurs il est alors possible de calculer également le pourcentage d'anthocyanes extractibles, ainsi que la maturité des pépins. Ce dosage a été effectué sur les 6 parcelles de raisins suivies en 2017, sur les 6 parcelles suivies en 2018, ainsi que sur les 15 parcelles suivies en 2019. Les résultats de chaque paramètre sont présentés par millésime sur la Figure 49. Sur cette représentation graphique de l'évolution des différents paramètres étudiés, la semaine 1 représente le premier prélèvement du millésime, et la dernière semaine (semaine 4 pour 2017 et 2018, semaine 5 pour 2019) représente le jour de vendanges. Les différentes parcelles suivies sont numérotées de M1 à M23 pour les Merlots, et de Cf1 à Cf4 pour les Cabernets francs.

Globalement, les trois paramètres étudiés montrent une évolution très influencée par l'évolution de la maturité. Nous constatons des teneurs variées en fonction des parcelles étudiées, mais celle-ci suivent une évolution commune durant la maturité. L'exemple le plus marqué pour illustrer ce constat est ApH1 durant 2017, où nous observons tout d'abord une augmentation des concentrations entre la semaine 1 et 2, suivie d'une diminution en semaine 3 pour finir avec une nouvelle augmentation pour la vendange. Seule la parcelle Cf1 montre une évolution différente en semaine 2. Concernant l'évolution des trois différents paramètres en fonction du millésime, nous constatons des évolutions divergentes. Ainsi, si l'on s'intéresse au millésime 2017, nous observons une évolution relativement proche entre ApH1 et la RPT, et une évolution propre à ApH3,2. A contrario, durant le millésime 2019, ce sont les paramètres ApH1 et ApH3,2 qui montrent des évolutions similaires, et la RPT qui évolue différemment. Ainsi, les évolutions sont différentes suivant les millésimes, et les paramètres ne montrent pas de corrélation entre eux. Il est dès lors intéressant de modéliser ces paramètres, car chacun d'entre eux apporte une information complémentaire sur l'évolution des polyphénols du raisin.

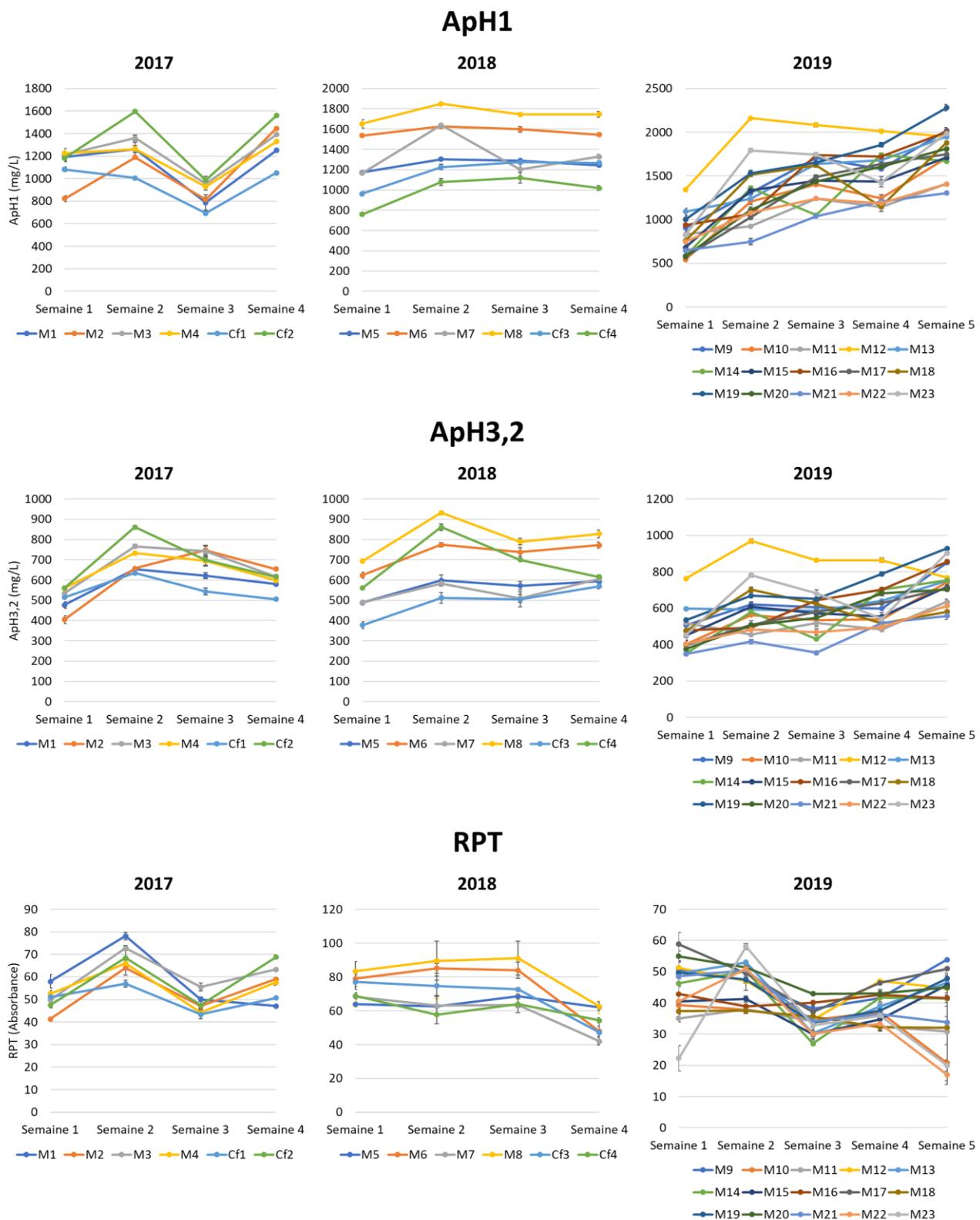


Figure 49 : Evolution du potentiel total en anthocyanes (ApH1), du potentiel en anthocyanes extractibles (ApH3,2), et de la richesse polyphénolique totale (RPT) sur les millésimes 2017, 2018 et 2019 (M : Merlot, Cf : Cabernet franc)

Si l'évolution de la maturité est un facteur important dans la variabilité étudiée, nous constatons également des différences de concentration en fonction de la parcelle. Au niveau des concentrations du potentiel en anthocyanes totales (ApH1), un écart notable est constaté en fonction de la parcelle étudiée, sur tous les millésimes. En s'intéressant au dernier prélèvement, le minimum pour les millésimes 2017, 2018 et 2018 est respectivement de 1050 mg/L (Cf1), 1242 mg/L (M5) et de 1304 mg/L (M21) et le maximum est respectivement de 1559 mg/L (Cf2), 1744 mg/L (M8) et de 2279 mg/L (M19). Ce type de variabilité est observable sur tous les paramètres, et pour tous les prélèvements. De plus, la plupart des parcelles étudiées montrent des évolutions qui leur sont propres. Par exemple, si l'on s'intéresse à l'évolution d'ApH3,2 de la parcelle M2 en 2017, nous remarquons qu'elle possède la concentration la plus faible pour le premier prélèvement, et la concentration la plus forte pour le dernier prélèvement. On remarque également une évolution divergente des autres parcelles pour M12 en 2019, pour les paramètres ApH1 et ApH3,2.

Ainsi, le jeu de données montre une variabilité importante sur les paramètres de la méthode Glories, avec des évolutions distinctes entre millésimes, paramètres et parcelles. Ces résultats reflètent donc la variabilité de la maturité du raisin, et apparaissent tout à fait adaptés pour la conception de modèles de prédiction.

## *(2) Conception de modèles de prédiction PLS pour les paramètres de la méthode Glories pendant la maturité du raisin*

En utilisant la régression PLS, des modèles de prédiction ont été construits pour l'UV-visible, l'IRTF, et la combinaison des deux zones spectrales. L'intérêt de cette combinaison est de mettre en évidence la possibilité ou non d'améliorer la prédiction en utilisant les données complémentaires des deux spectres. Le Tableau 13 présente les résultats de prédictions obtenus en utilisant tout le jeu de données. Les différents paramètres des modèles sont :

- $R^2C$  : Coefficient de régression linéaire de la calibration, qui permet de juger de la robustesse de la calibration.
- RMSEC : Erreur quadratique de calibration.
- $R^2CV$  : Coefficient de régression linéaire de la validation croisée. Cette donnée va être un premier indicateur de la robustesse de prédiction du modèle.
- RMSECV : Erreur quadratique de validation croisée

- RPD CV : Ecart prédictif résiduel de la validation croisée. Il est calculé par le rapport entre l'erreur quadratique de validation croisée et l'écart-type du paramètre étudié. Lorsque le RPD est supérieur à 2, le modèle peut être considéré comme bon, et lorsqu'il est supérieur à 3, il est considéré comme excellent (6,10,270).
- VL : Nombre de variables latentes utilisées pour la conception du modèle. Ce choix est effectué pour minimiser la valeur de RMSECV.

Pour tous les modèles PLS présentés, tous les triplicats d'analyse sont utilisés. Cela permet d'intégrer dans le modèle l'erreur induite par les analyses de références et les analyses spectrales. Si cette configuration peut diminuer les résultats de prédiction, cela permet d'obtenir des modèles qui reflètent mieux la réalité des analyses. Le prétraitement des données est un centrage par la moyenne. Les paramètres de validation croisée sont les suivants : 10 groupes sont définis par une répartition de type *venetian blinds*. La répartition est faite en tenant compte des réplicats, pour éviter une surestimation de la validation croisée.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL
<b>UV-visible</b>						
ApH1 (mg/L)	0,84	152,28	0,74	196,87	1,95	11
ApH3,2 (mg/L)	0,77	67,41	0,59	90,12	1,55	11
<b>IRTF</b>						
ApH1 (mg/L)	0,88	130,03	0,81	169,82	2,26	13
ApH3,2 (mg/L)	0,896	44,86	0,83	56,47	2,47	13
RPT (absorbance)	0,73	8,05	0,63	9,6	1,63	13
<b>IRTF-UV-visible</b>						
ApH1 (mg/L)	0,88	131,79	0,81	167,16	2,30	12
ApH3,2 (mg/L)	0,82	59,25	0,70	76,15	1,83	12

Tableau 13 : Modèles de prédiction PLS des paramètres de la méthode Glories par UV-visible, IRTF et IRTF-UV-visible. Utilisation du jeu de données complet pour la calibration (RPT : richesse polyphénolique totale, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Compte tenu du fait que les analyses spectrales ont été effectuées sur le résultat de l'extraction à pH 3,2, les prédictions présentées pour ApH1 sont induites indirectement. Le but est donc de rechercher s'il est possible de corrélérer les informations apportées par les spectres à pH 3,2 et les valeurs d'ApH1.

Les modèles de prédiction d'ApH1 pour l'UV-visible, l'IRTF et la combinaison des deux techniques montrent des résultats très proches. Pour l'IRTF et l'IRTF-UV-visible, le  $R^2C$  et le RMSEC sont pratiquement identiques, avec respectivement 0,88 et 130,03 mg/L pour l'IRTF, et 0,88 et 131,79 mg/L pour l'IRTF-UV-vis. Les résultats obtenus pour l'UV-visible sont légèrement moins performants, avec 0,84 et 152,28 mg/L. Les résultats de calibration sont un premier indicateur pour juger la possibilité ou non de concevoir un modèle de prédiction, mais ne permettent pas de définir la robustesse de prédiction. Analyser ces paramètres seuls n'est pas pertinent, et ils doivent être comparés à ceux de la validation. La littérature n'indique pas à partir de quelle valeur il est possible de considérer une calibration ou une validation efficace, car cela dépend notamment de l'utilisation ultérieure du modèle et si l'erreur est acceptable ou non par l'utilisateur. Dans notre cas, nous considérerons que si le coefficient de régression de validation est inférieur à 0,8, la linéarité de la prédiction n'est plus acceptable. Cependant, ce résultat seul n'est pas suffisant il doit être associé à l'erreur de validation. En effet, une prédiction peut montrer des résultats linéaires mais déviants de la calibration. Le paramètre le plus pertinent est alors le RPD, car il permet de comparer l'erreur de prédiction à l'écart type du paramètre. Si l'on s'intéresse aux résultats de validation croisée d'ApH1, nous constatons que le modèle produit avec les spectres UV-visible ne montre pas une robustesse suffisante. Le RPD CV est légèrement inférieur à 2 (1,95) et le  $R^2CV$  est de 0,74. L'UV-visible ne semble donc pas adapté à la prédiction du potentiel total en anthocyanes. Si l'on s'intéresse maintenant aux résultats de l'IRTF et de la combinaison nous constatons des prédictions plus efficaces. Pour l'IRTF, le  $R^2CV$  est de 0,81 et le RPD CV de 2,26. Pour l'IRTF-UV-visible, le  $R^2CV$  est de 0,81 et le RPD CV de 2,30. Étonnamment, si l'UV-visible ne paraît pas efficace pour la prédiction d'ApH1, le résultat de sa combinaison avec l'IRTF montre des résultats très légèrement supérieurs à l'IRTF seul. Si l'on s'intéresse à l'erreur de validation croisée, nous constatons cependant que la différence entre l'IRTF et la combinaison est très faible, inférieure à 3 mg/L. Il est donc pour le moment impossible de juger de la pertinence ou non de cette combinaison. Dans cet exemple, le gain de robustesse apparaît trop faible en comparaison de l'instrumentation demandée.

Les résultats pour ApH3,2 montrent des divergences bien plus importantes que pour ApH1. Pour l'UV-visible, les résultats de la validation croisée sont faibles, avec un  $R^2CV$  de 0,59 et un RPD CV de 1,55. Étonnamment, ces résultats sont plus faibles que pour ApH1, alors que l'enregistrement du spectre UV-visible a été fait sur la solution à pH 3,2. Ces valeurs semblent valider le fait que l'UV-visible n'est pas adapté pour la prédiction de la méthode Glories. Compte tenu de la manipulation nécessaire pour obtenir ApH3,2, à savoir une différence d'absorbance à 520 nm entre une solution

témoin et une solution décolorée au bisulfite de sodium, il est surprenant de ne pas observer une corrélation avec l'UV-visible plus forte. Cependant, afin d'obtenir le spectre nous avons effectué une dilution au 1/100, qui influence la copigmentation des anthocyanes (342). Une hypothèse envisageable serait que cette manipulation ait pu influencer la prédiction. Les résultats de la prédiction avec l'IRTF montrent quant à eux une robustesse supérieure aux résultats d'ApH1, avec un  $R^2CV$  de 0,83 et un RPD CV de 2,47. Ainsi, l'IRTF apparaît comme une solution efficace pour la prédiction de la méthode Glories. Concernant les l'IRTF-UV-visible, l'erreur de prédiction attribuée à l'UV-visible semble fortement impacter les résultats, qui se montrent inférieurs à ceux de l'IRTF seul, avec un  $R^2CV$  de 0,70 et un RPD CV de 1,83. Le gain de robustesse observé pour ApH1 n'est pas confirmé pour ApH3,2.

Le dernier paramètre étudié est la richesse polyphénoliques totale (RPT). Comme il s'agit de la lecture directe d'une absorbance à 280 nm, les essais de prédiction n'ont porté que sur l'IRTF. Les résultats obtenus, à savoir un  $R^2CV$  de 0,63 et un RPD CV de 1,63 montrent clairement l'inaptitude de l'IRTF à prédire correctement ce paramètre.

Ces premiers modèles indiquent des résultats mitigés pour la prédiction de la méthode Glories pendant la maturité du raisin. Si l'UV-visible ne se montre pas efficace, les résultats de l'IRTF sont encourageants pour les paramètres ApH1 et ApH3,2. Cependant, nous avons constaté précédemment un fort effet millésime sur le jeu de données. Il est donc nécessaire de s'assurer de la prédiction d'une année à l'autre.

### *(3) Essai de validation des modèles de prédiction PLS en fonction du millésime pendant la maturité du raisin*

Pour valider les prédictions des modèles obtenus précédemment, le jeu de données a été divisé en un groupe de calibration et un groupe de validation. Le groupe de validation est composé des échantillons d'un millésime, et le groupe de validation des deux millésimes restants. Ainsi, nous avons construit des modèles avec une calibration des millésimes 2018 et 2019 avec 2017 pour la validation (Cal 2018-2019, Val 2017), et avec une calibration des millésimes 2017 et 2019 avec 2018 pour la validation (Cal 2017-2019, Val 2018). Le millésime 2019 n'a pas été utilisé pour la validation, car il compose plus de la moitié du jeu de données et n'aurait pas permis de laisser assez d'échantillons pour une calibration correcte. Les résultats de prédiction sont présentés dans le Tableau 14.

Paramètre	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL
	Cal 2018-2019, Val 2017				Cal 2017-2019, Val 2018			
<b>UV-visible</b>								
ApH1 (mg/L)	0,09	439,77	0,87	9	0,34	236,09	1,63	9
ApH3,2 (mg/L)	0,28	116,55	1,2	9	0,45	112,1	1,24	10
<b>IRTF</b>								
ApH1 (mg/L)	0,23	236,45	1,62	3	0,62	183,65	2,09	3
ApH3,2 (mg/L)	0,66	64,23	2,17	13	0,85	119,2	1,17	11
RPT (absorbance)	0,06	10,27	1,52	6	0,07	17,82	0,88	7
<b>IRTF-UV-visible</b>								
ApH1 (mg/L)	0,21	318,77	1,21	9	0,57	208,17	1,85	8
ApH3,2 (mg/L)	0,53	83,83	1,66	7	0,67	96,63	1,44	10

Tableau 14 : Résultats de prédiction des modèles PLS des paramètres de la méthode Glories en fonction du millésime (RPT : richesse polyphénolique totale R<sup>2</sup>P : coefficient de régression de la prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Tous les résultats présentés indiquent des valeurs de prédiction très faibles en comparaison aux valeurs obtenus sur le jeu de données complet. L'effet millésime apparaît donc être un facteur limitant la robustesse des modèles. Nous remarquons également que les coefficients de régression de prédiction R<sup>2</sup>P sont tous plus faibles pour la validation avec 2017 qu'avec 2018. La variabilité de 2017 semble donc plus dure à prédire que 2018.

Les résultats de l'UV-visible et de l'IRTF-UV-visible sont peu prédictifs, avec des R<sup>2</sup>P inférieurs à 0,7 et des RPD P inférieurs à 2. Ces données confirment les observations faites sur le jeu de données complet. Intéressons-nous à l'IRTF, qui avait montré le meilleur potentiel pour la prédiction d'ApH1 et ApH3,2. Les valeurs de prédiction de la richesse polyphénolique totale (R<sup>2</sup>P < 0,1) montrent l'inaptitude de l'IRTF à prédire ce paramètre et ne seront pas discutées. Concernant ApH1, nous constatons des résultats très différents suivant le millésime de validation. Le R<sup>2</sup>P varie de 0,23 pour la prédiction de 2017 à 0,62 pour 2018. Concernant le RPD P, il varie de 1,62 pour 2017 à 2,09 pour 2018. Si les résultats sont insuffisants pour la validation avec 2017, nous remarquons que malgré un R<sup>2</sup>P faible, l'erreur de prédiction du millésime 2018 apparaît acceptable. Cette différence entre la linéarité de la prédiction et son erreur est encore plus marquée pour ApH3,2. En effet, pour la prédiction du millésime 2017, le R<sup>2</sup>P peut être considéré comme faible (0,66), alors que le RPD P est acceptable (2,17). A l'inverse, pour la prédiction de 2018, le R<sup>2</sup>P est acceptable (0,85) et le RPD P faible (1,17). Afin de comprendre l'origine de cette divergence, les valeurs d'ApH3,2 mesurées en fonction des valeurs prédites par les modèles sont présentées en Figure 50.

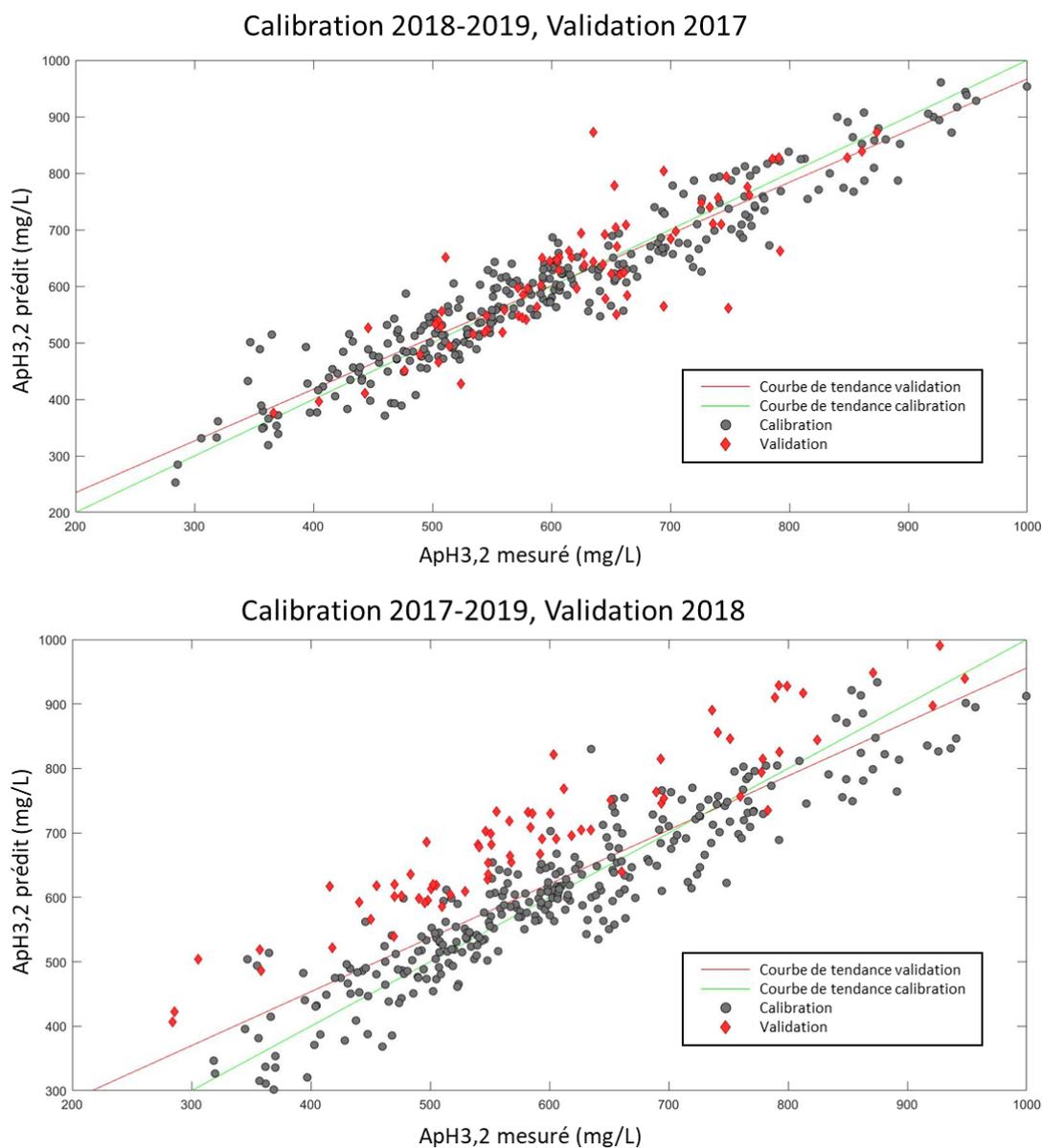


Figure 50 : ApH3,2 mesuré en fonction d'ApH3,2 prédit par les modèles de prédiction PLS de l'IRTF. Séparation de la calibration et de la validation en fonction des millésimes.

Pour la validation de 2017, les échantillons de validation s'intègrent à la linéarité de la calibration, à l'exception de plusieurs échantillons qui s'écartent fortement. Ces données vont influencer la linéarité de la prédiction, et donc le  $R^2P$ . Cependant, comme la majorité des échantillons montrent une prédiction proche de la calibration, les valeurs déviantes ne vont que très peu impacter l'erreur moyenne de prédiction. Le modèle présenté est donc efficace, mais montre une inaptitude à prédire correctement certains échantillons.

Pour la validation de 2018, nous constatons que les valeurs de prédiction sont linéaires. Cependant, tous les échantillons de la validation montrent des concentrations prédites surestimées. Cette erreur systématique explique alors l'erreur de prédiction moyenne importante.

Dans l'ensemble, l'UV-visible apparaît inadapté pour la prédiction de la méthode Glories sur le jeu de données étudié. L'IRTF montre des résultats encourageants sur le paramètre ApH3,2, mais pas suffisants pour ApH1 et RPT. L'intérêt de la méthode Glories repose sur l'analyse combinée de ces trois paramètres, l'IRTF n'est donc pas assez performant. Cependant, les résultats obtenus avec l'IRTF semblent indiquer la possibilité de prédire le potentiel en anthocyane extractible, et doivent être approfondis avec des mesures plus précises, notamment le dosage des anthocyanes moléculaires par CLHP.

### ***b) Dosage des anthocyanes moléculaires sur les millésimes 2018 et 2019 pendant la maturité du raisin***

#### *(1) Evolution des anthocyanes moléculaires pendant la maturité du raisin*

Les concentrations en anthocyanes moléculaires ont été obtenues par dosage en chromatographie liquide haute performance et détection UV-visible (CLHP-UV-visible). Cette mesure a été effectuée sur les 6 parcelles de raisins suivies en 2018, ainsi que sur les 15 parcelles suivies en 2019. Les résultats de la somme des anthocyanes moléculaires sont présentés par millésime sur la Figure 51. La numérotation des parcelles suivies est identique à celle utilisée pour la méthode Glories.

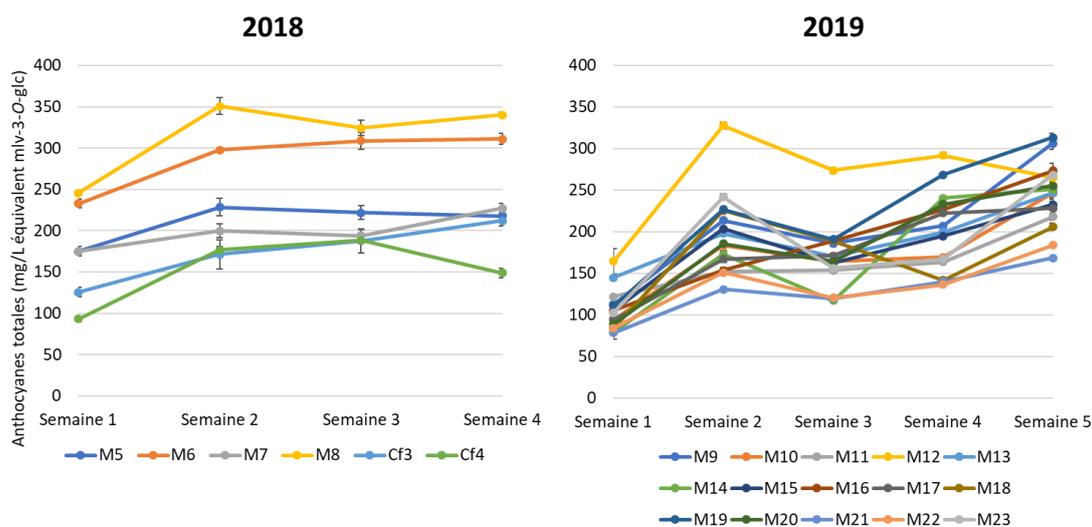


Figure 51 : Evolution de la concentration en anthocyanes moléculaires sur les millésimes 2018 et 2019 (M : Merlot, Cf : Cabernet franc)

Nous constatons une évolution des anthocyanes moléculaires fortement corrélée à l'évolution d'ApH3,2 présentée en Figure 49, avec quelques différences observables. La plus visible est l'évolution de la parcelle Cf4, qui indique des concentrations plus élevées pour ApH3,2 en comparaison des autres parcelles. La méthode Glories est moins précise et moins spécifique que le dosage par CLHP-UV-visible, ce qui peut expliquer ces différences.

Tout comme ApH3,2 discuté précédemment, le dosage des anthocyanes moléculaires montre une variabilité importante avec des évolutions distinctes entre millésimes, paramètres et parcelles, et est donc adapté à la conception de modèles de prédiction. Nous avons pu constater que les différentes concentrations d'anthocyanes sont corrélées avec la concentration en anthocyanes totales lors de l'ACP présentée en Figure 43. Il est donc nécessaire d'étudier la variabilité du ratio des différentes anthocyanes entre-elles. La Figure 52 présente la distribution des anthocyanes en fonction des parcelles étudiées, exprimée en pourcentages. L'écart-type correspond à la variabilité de chaque anthocyanes pour tous les prélèvements d'une même parcelle.

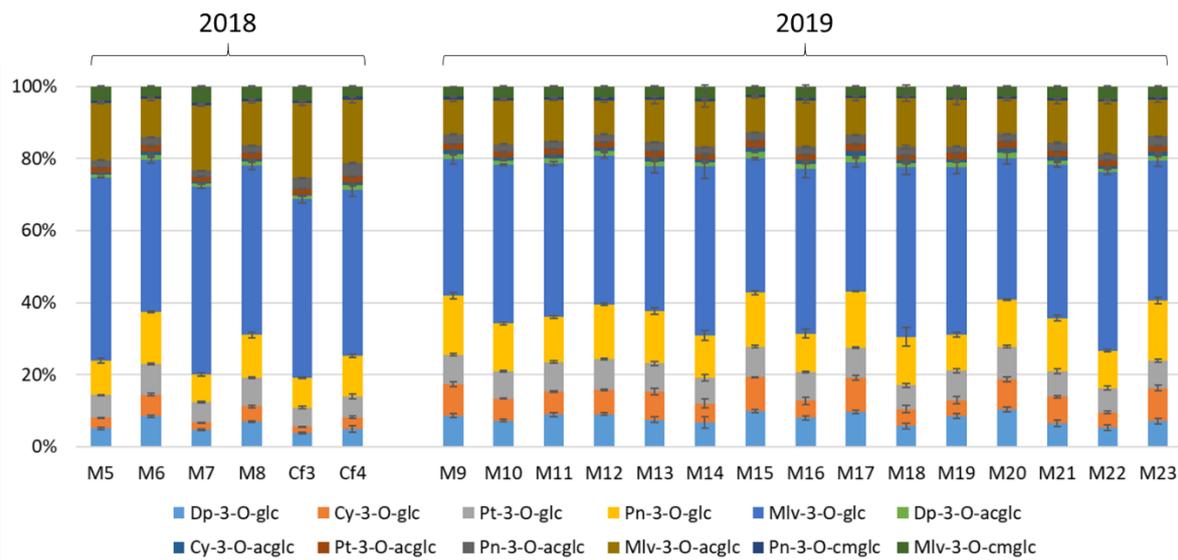


Figure 52 : Distribution des anthocyanes en fonction des parcelles étudiées (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside)

L'évolution de la composition en anthocyanes apparaît relativement similaire entre parcelles. Les seuls composés qui montrent une différence notable sont la malvidine-3-O-glucoside et la cyanidine-3-O-glucoside. Nous constatons que lorsque le pourcentage de malvidine-3-O-glucoside est supérieur à la tendance générale, celui de la cyanidine-3-O-glucoside est inférieur. Ces résultats sont en accord avec l'ACP des anthocyanes moléculaires (Figure 43) qui montrait la variable malvidine-3-O-glucoside inversement corrélée à la variable cyanidine-3-O-glucoside par rapport à la dimension 2. Si l'on s'intéresse à la variation du pourcentage de chaque anthocyane par parcelle, représentée par la barre d'erreur, nous constatons que celle-ci est très faible. La répartition des anthocyanes apparaît donc définie dès la première semaine d'échantillonnage, et n'évolue que très peu durant la maturité.

La variabilité de chaque anthocyane en fonction du total des anthocyanes apparaît peu importante lors de la comparaison des parcelles entre elles, ainsi que lors de l'évolution de la maturité. Différentes études sur les anthocyanes du raisin décrivent des résultats similaires (95,343,344).La concentration de chacun de ces composés sera donc corrélée avec la concentration en anthocyanes totales, comme nous avons pu le décrire lors de discussion des ACP des variables. Si les résultats de prédiction d'une anthocyane particulière montre une robustesse similaire à la prédiction des anthocyanes totales, il sera alors difficile de distinguer une corrélation propre au composé étudié, et non une corrélation indirecte avec les anthocyanes totales.

(2) *Conception et validation de modèles de prédiction PLS pour le dosage des anthocyanes moléculaires pendant la maturité du raisin*

En utilisant la régression PLS, des modèles de prédiction ont été construits pour l'IRTF et l'UV-visible. Chacune des 12 anthocyanes majoritaires du raisin a été étudiée, ainsi que la somme des anthocyanes-3-*O*-glucoside, des anthocyanes-3-*O*-acétylglucoside, des anthocyanes-3-*O*-coumaroylglucoside et des anthocyanes moléculaires. Compte tenu du nombre important de paramètres étudiés, les résultats seront divisés en fonction de l'analyse spectrale utilisée.

Modèles de prédiction PLS du dosage des anthocyanes moléculaires par IRTF pendant la maturité du raisin

Nous avons constaté avec les modèles de prédiction de la méthode Glories que les résultats de calibration et de validation croisée sur le jeu de données complet ne reflétaient pas nécessairement la réalité de la prédiction d'échantillons extérieurs. Les résultats de prédiction seront donc discutés en parallèle. Deux types de validation avec des échantillons extérieurs à la calibration ont été utilisés. En premier lieu, le jeu de données a été divisé en un set de calibration et de validation grâce à l'algorithme de Kennard-Stone (340). Cela permet d'obtenir un set de validation représentatif de la variabilité des échantillons. Pour la deuxième validation, les données de 2019 ont été utilisées pour la calibration, et celles de 2018 pour la validation. Afin de pouvoir comparer ces deux validations, le nombre d'échantillons pour la calibration avec l'algorithme de Kennard-Stone a été fixé à 75, ce qui correspond au nombre d'échantillons pour le millésime 2019. Les résultats obtenus pour l'IRTF sont présentés dans le Tableau 15.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL
Type de validation	Validation croisée						Cal/Val Kennard-Stone				Cal 2019, Val 2018			
Dp-3-O-glc (mg/L)	0,79	3,21	0,68	4,06	1,74	11	0,62	4,32	1,63	10	0,57	7,11	0,99	9
Cy-3-O-glc (mg/L)	0,81	2,66	0,68	3,54	1,73	11	0,75	2,99	2,05	7	0,37	6,16	1,00	8
Pt-3-O-glc (mg/L)	0,86	2,25	0,78	2,88	2,11	11	0,80	2,82	2,16	11	0,72	5,60	1,09	9
Pn-3-O-glc (mg/L)	0,85	4,01	0,74	5,32	1,94	11	0,83	4,40	2,34	6	0,53	11,54	0,89	9
Mlv-3-O-glc (mg/L)	0,93	8,08	0,87	11,11	2,74	12	0,87	11,41	2,67	12	0,82	14,22	2,14	8
Dp-3-O-acglc (mg/L)	0,77	0,53	0,60	0,71	1,56	12	0,68	0,62	1,79	11	0,52	1,25	0,89	9
Cy-3-O-acglc (mg/L)	0,79	0,33	0,66	0,44	1,68	11	0,72	0,39	1,90	11	0,39	0,86	0,86	9
Pt-3-O-acglc (mg/L)	0,86	0,45	0,79	0,57	2,16	11	0,78	0,64	1,92	9	0,72	0,74	1,66	9
Pn-3-O-acglc (mg/L)	0,74	0,77	0,65	0,65	2,34	8	0,71	0,85	1,79	7	0,22	1,38	1,10	9
Mlv-3-O-acglc (mg/L)	0,84	3,86	0,79	4,54	2,15	8	0,89	3,35	2,92	13	0,43	7,12	1,37	10
Pn-3-O-cmglc (mg/L)	0,67	0,32	0,60	0,36	1,56	7	0,73	0,27	2,07	7	0,50	0,48	1,17	9
Mlv-3-O-cmglc (mg/L)	0,89	0,89	0,81	1,18	2,28	12	0,83	1,15	2,34	12	0,56	1,77	1,52	10
Total glucoside (mg/L)	0,93	13,46	0,88	17,86	2,89	12	0,77	23,50	2,19	7	0,77	33,31	1,55	9
Total acétyl (mg/L)	0,92	3,51	0,86	4,59	2,68	12	0,89	4,26	2,89	12	0,60	6,35	1,94	7
Total coumaroyl (mg/L)	0,90	0,96	0,83	1,26	2,43	12	0,86	1,23	2,49	12	0,64	1,78	1,72	10
Total anthocyanes (mg/L)	0,94	16,37	0,90	20,63	3,16	12	0,91	19,52	3,33	12	0,80	32,68	1,99	9

Tableau 15 : Modèles de prédiction PLS du dosage des anthocyanes moléculaires par IRTF (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Pour les résultats de validation croisée, l'écart prédictif résiduel (RPD CV) montre des valeurs comprises entre 3,16 pour les anthocyanes totales et 1,56 pour la péonidine-3-*O*-coumaroylglucoside et la delphinidine-3-*O*-acétylglucoside. S'il n'est pas surprenant de constater que la concentration en anthocyanes totales est la mieux modélisée, les données de validation croisée concernant les anthocyanes minoritaires montrent une robustesse qui n'était pas forcément attendue. Si l'on s'intéresse à l'erreur de validation croisée, nous constatons que pour plusieurs anthocyanes celle-ci est inférieure à 1 mg/L, ce qui apparaît extrêmement faible en comparaison à l'erreur des anthocyanes totales, égale à 20,63 mg/L. Intéressons-nous à la modélisation de la pétunidine-3-*O*-acétylglucoside, car les résultats de prédiction de cette anthocyane apparaissent élevés ( $R^2_{CV} : 0,79$ , RPD CV : 2,16) en comparaison à sa concentration moyenne (3,29 mg/L). L'erreur de prédiction du modèle est de 0,57 mg/L, alors que l'erreur de l'analyse de référence est de 0,23 mg/L. Ces résultats impliquent un modèle avec une performance largement inférieure au mg/L, proche de l'erreur d'analyse CLHP-UV-visible, et très éloignée des valeurs auxquelles on pourrait s'attendre pour la spécificité du composé étudié dans une matrice complexe. L'hypothèse la plus probable pour expliquer cette prédiction est que la modélisation de la pétunidine-3-*O*-acétylglucoside ne s'appuie pas sur l'empreinte spectrale du composé, mais est corrélée indirectement avec d'autres valeurs. Nous avons constaté lors de l'étude de l'ACP des anthocyanes moléculaires que la pétunidine-3-*O*-acétylglucoside était fortement corrélée avec la concentration en anthocyanes totales (Figure 43) et pourrait expliquer les résultats. Afin de valider cette hypothèse, nous avons tenté de prédire les concentrations en pétunidine-3-*O*-acétylglucoside grâce au modèle construit pour les anthocyanes totales. La représentation graphique de la prédiction du modèle est présentée en Figure 53.

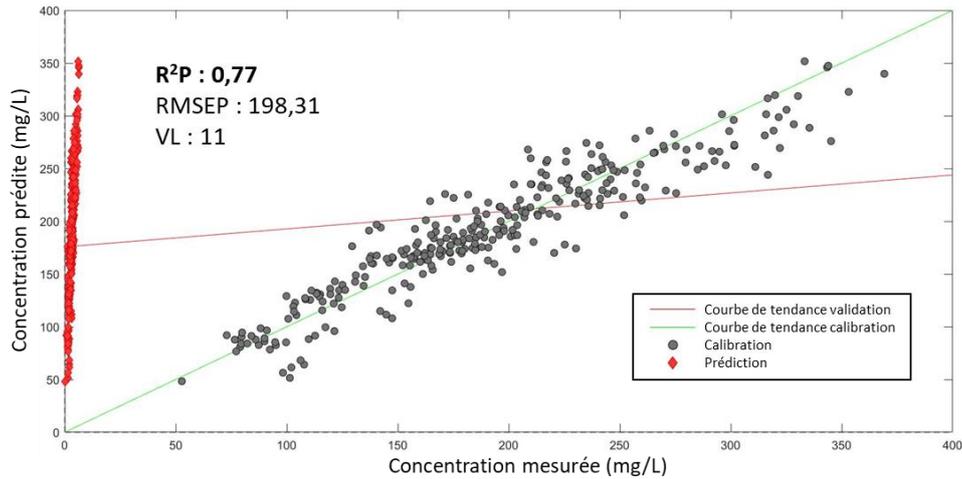


Figure 53 : Prédiction de la concentration en pétunidine-3-*O*-acétylglucoside en fonction du modèle de prédiction des anthocyanes totales par IRTF ( $R^2P$  : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, VL : nombre de variables latentes)

Les résultats obtenus pour ce modèle montrent un  $R^2P$  de 0,78, équivalent au  $R^2CV$  du modèle de la pétunidine-3-*O*-acétylglucoside (0,79). La linéarité de la prédiction, observable sur la Figure 53, est donc identique pour les deux modèles. L'erreur de prédiction est importante car le modèle des anthocyanes totales n'est pas conçu pour la prédiction de la concentration en pétunidine-3-*O*-acétylglucoside. Il est cependant possible de rechercher le coefficient de régression linéaire de la corrélation entre les concentrations en pétunidine-3-*O*-acétylglucoside et en anthocyanes totales (Figure 54).

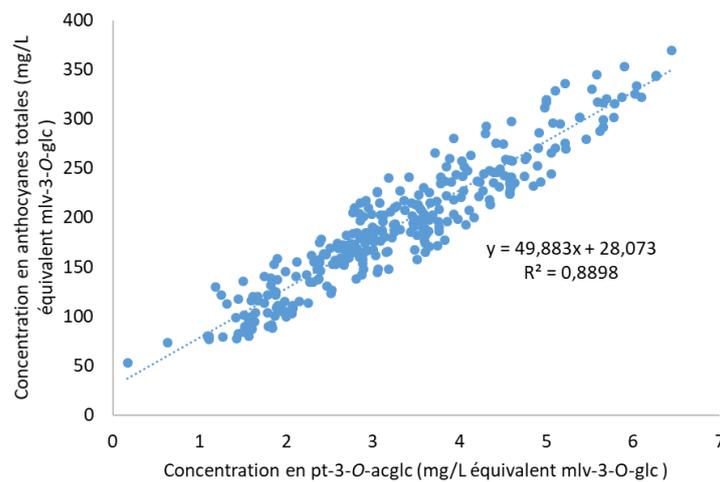


Figure 54 : Corrélation entre les concentrations en pétunidine-3-*O*-acétylglucoside et en anthocyanes totales

En appliquant l'équation de la droite de corrélation aux concentrations en pétunidine-3-*O*-acétylglucoside, nous obtenons une concentration modifiée que nous pouvons de nouveau essayer de prédire avec le modèle de prédiction conçu pour les anthocyanes totales (Figure 55)

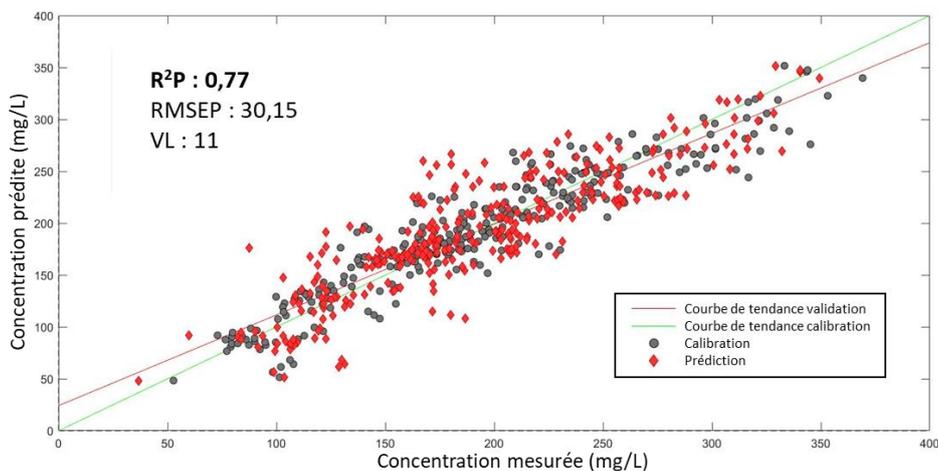


Figure 55 : Prédiction de la concentration modifiée de pétunidine-3-*O*-acétylglucoside en fonction du modèle de prédiction des anthocyanes totales par IRTF ( $R^2P$  : coefficient de régression de prédiction,  $RMSEP$  : erreur quadratique moyenne de prédiction,  $VL$  : nombre de variables latentes)

Nous constatons que la prédiction est identique à celle obtenue précédemment, avec cette fois ci une distribution de la prédiction proche de la calibration. L'erreur de prédiction est bien inférieure (30,15 mg/L), et une fois divisée par le coefficient directeur appliqué aux données (49,88), nous obtenons une erreur de prédiction égale à 0,60 mg/L, très proche de l'erreur de validation croisée de 0,57 mg/L obtenue pour le modèle de la pétunidine-3-*O*-acétylglucoside. Il est donc démontré que la prédiction de la pétunidine-3-*O*-acétylglucoside est une corrélation indirecte avec les concentrations en anthocyanes totales. Le même raisonnement peut être appliqué à toutes les autres anthocyanes mesurées, et nous ne pouvons pas attribuer de prédiction pour celles-ci.

Dans la littérature, plusieurs études sur le raisin, moût et vin utilisant l'analyse multivariée couplée à l'IRTF ou l'UV-visible ont démontré des capacités à prédire spécifiquement chaque anthocyane moléculaire (12,13,287,329). Pour chacun de ces travaux, les erreurs de prédiction ou de calibration des concentrations en pétunidine-3-*O*-acétylglucoside rapportées sont inférieures au mg/L, et apparaissent fortement corrélées avec celles obtenues pour les anthocyanes totales. Nous pouvons alors avancer que les résultats présentés ne sont pas dus à une modélisation spécifique de chaque anthocyane moléculaire, mais à une corrélation indirecte avec la concentration en anthocyanes totales. Cette erreur de corrélation n'a jusqu'à présent jamais été reportée ou discutée dans la littérature pour la prédiction des anthocyanes. Nous constatons donc les dérives possibles de l'utilisation de l'analyse

multivariée sur une matrice complexe comme le vin, où les corrélations entre composés peuvent amener à surestimer les modèles obtenus. Si les résultats présentés dans le Tableau 15 laissent à penser que la prédiction de chaque anthocyane est envisageable, nous ne pouvons pas considérer ces résultats comme fiables. La validité ou non d'un modèle de prédiction dépend avant tout des besoins de l'utilisateur, et il est tout à fait possible d'utiliser une corrélation indirecte. Cependant, si le modèle recherché se veut représentatif de l'empreinte spectrale du composé étudié, alors seul le résultat obtenu pour les anthocyanes totales doit être pris en compte.

En s'intéressant à ces résultats, nous constatons des valeurs de prédictions par IRTF pour la modélisation de la somme des anthocyanes moléculaires robustes. Les données de calibration ( $R^2C$  : 0,94, RMSEC : 16,37 mg/L) et de validation croisée ( $R^2C$  : 0,90, RMSEC : 20,63 mg/L) sont très proches, et indiquent une validation croisée pratiquement aussi performante que la calibration. Le RPD CV de 3,16 indique un modèle performant qui peut être utilisé pour la prédiction des anthocyanes du raisin pendant la maturité (10). Cependant, nous avons constaté avec la prédiction de la méthode Glories que la validation croisée n'est pas forcément suffisante pour juger de l'efficacité d'un modèle, et doit être confrontée à la prédiction d'échantillons extérieurs. Les résultats de prédiction obtenus en divisant le jeu de données grâce à l'algorithme de Kennard-Stone confirme la validation croisée, avec un  $R^2P$  de 0,91, une erreur de prédiction (RMSEP) de 19,52 mg/L et un RPD P de 3,33. Les valeurs de prédiction apparaissent même supérieures à celles de la validation croisée. Compte tenu du fort effet millésime mis en évidence précédemment, les données obtenues doivent être confrontées à la prédiction d'échantillons d'un millésime extérieurs à la calibration. Les résultats obtenus avec comme calibration le millésime 2019 et comme validation le millésime 2018 sont de 0,80 pour le  $R^2P$ , de 32,68 mg/L pour le RMSEP et de 1,99 pour le RPD P. Ces valeurs indiquent une diminution de la prédiction, mais restent très inférieures à celles observées pour la méthode Glories. De plus, avec un RPD très proche de 2, les résultats obtenus montrent une erreur considérée comme acceptable (10).

Cette étude des anthocyanes moléculaires par IRTF nous a permis de mettre en évidence les erreurs de prédiction possibles. Si l'analyse multivariée est un outil puissant, elle peut aussi amener à mal interpréter les résultats obtenus. De plus, nous avons pu avancer des résultats de prédiction robustes pour les anthocyanes totales, qui démontrent la faculté de l'IRTF à prédire ces composés. Si l'effet millésime impacte la prédiction, le modèle présenté reste cependant fiable pour être utilisé.

## Modèles de prédiction PLS du dosage des anthocyanes moléculaires par UV-visible pendant la maturité du raisin

Les valeurs des modèles de prédiction des anthocyanes moléculaires par UV-visible sont reportées dans le Tableau 16. Afin de pouvoir comparer ces résultats avec ceux obtenus pour l'IRTF, toutes les anthocyanes moléculaires sont présentées, avec les données obtenues pour la calibration et la validation croisée, ainsi que pour la validation utilisant l'algorithme de Kennard-Stone et la validation par millésime.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL
Type de validation	Validation croisée						Cal/Val Kennard-Stone				Cal 2019, Val 2018			
Dp-3-O-glc (mg/L)	0,82	2,99	0,68	3,99	1,77	11	0,60	3,91	1,81	10	0,07	74,04	0,10	9
Cy-3-O-glc (mg/L)	0,80	2,74	0,65	3,61	1,70	11	0,54	3,97	1,55	11	0,08	80,96	0,08	12
Pt-3-O-glc (mg/L)	0,84	2,42	0,71	3,25	1,87	11	0,64	3,04	2,00	10	0,05	57,59	0,11	9
Pn-3-O-glc (mg/L)	0,81	4,46	0,67	5,91	1,74	11	0,56	6,26	1,65	10	0,03	81,25	0,13	12
Mlv-3-O-glc (mg/L)	0,87	11,01	0,80	13,57	2,24	10	0,81	11,23	2,71	10	0,12	98,09	0,31	9
Dp-3-O-acglc (mg/L)	0,74	0,56	0,61	0,69	1,61	10	0,60	0,64	1,73	10	0,07	12,87	0,09	9
Cy-3-O-acglc (mg/L)	0,73	0,38	0,59	0,48	1,54	10	0,44	0,53	1,40	10	0,11	10,39	0,07	12
Pt-3-O-acglc (mg/L)	0,80	0,56	0,69	0,68	1,81	10	0,68	0,64	1,92	10	0,05	12,74	0,10	9
Pn-3-O-acglc (mg/L)	0,77	0,73	0,64	0,92	1,65	10	0,54	0,90	1,69	10	0,03	12,75	0,12	12
Mlv-3-O-acglc (mg/L)	0,86	3,69	0,79	4,48	2,18	9	0,84	3,19	3,07	10	0,17	17,74	0,55	8
Pn-3-O-cmglc (mg/L)	0,71	0,30	0,51	0,39	1,44	10	0,38	0,36	1,56	10	0,06	3,95	0,14	12
Mlv-3-O-cmglc (mg/L)	0,84	1,09	0,75	1,36	1,98	10	0,79	1,04	2,59	10	0,08	6,53	0,41	8
Total glucoside (mg/L)	0,85	20,01	0,77	24,73	2,09	10	0,74	22,27	2,32	10	0,05	432,58	0,12	9
Total acétyl (mg/L)	0,86	4,53	0,79	5,64	2,18	10	0,81	6,71	1,83	10	0,06	62,98	0,20	9
Total coumaroyl (mg/L)	0,83	1,26	0,74	1,57	1,95	10	0,77	1,22	2,51	10	0,05	11,56	0,26	8
Total anthocyanes (mg/L)	0,86	24,62	0,78	30,21	2,15	10	0,77	27,34	2,38	10	0,05	505,38	0,13	9

Tableau 16 : Modèles de prédiction PLS du dosage des anthocyanes moléculaires par UV-visible (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Les résultats obtenus pour la modélisation des anthocyanes moléculaires par UV-visible apparaissent similaires à ceux de l'IRTF. Tout comme pour les modèles avec l'IRTF, nous constatons pour les anthocyanes minoritaires du raisin des erreurs de calibration et de validation croisée inférieures au mg/L. Ces résultats semblent donc indiquer comme démontré précédemment, que la modélisation des anthocyanes moléculaires est due à une corrélation avec les concentrations en anthocyanes totales. Cette hypothèse est corroborée par le fait que le nombre de variables latentes optimales (VL) nécessaires à la conception des modèles est pratiquement identique pour tous les paramètres. La variabilité des différentes concentrations étudiées sur les spectres UV-visible est donc expliquée par le même nombre de variables. Afin de valider cette hypothèse, nous nous sommes intéressés à la prédiction de la pétunidine-3-*O*-acétylglucoside à l'aide du modèle de prédiction des anthocyanes totales. Le R<sup>2</sup>P obtenu est de 0,72, pratiquement identique au R<sup>2</sup>CV de 0,69 observé pour le modèle de la pétunidine-3-*O*-acétylglucoside. Ainsi, tout comme pour l'IRTF, nous pouvons avancer que la prédiction des différentes anthocyanes moléculaires est due à leur corrélation avec les concentrations en anthocyanes totales, et donc à une modélisation indirecte qui ne correspond pas au composé étudié. De ce fait, seuls les résultats des anthocyanes totales peuvent être considérés et seront discutés.

La modélisation des anthocyanes totales montrent des résultats exploitables pour la validation croisée (R<sup>2</sup>CV : 0,78, RPD CV : 2,15) et la validation avec l'algorithme de Kennard-Stone (R<sup>2</sup>P : 0,77, RPD P : 2,38). Ces valeurs sont plus faibles que celles obtenues pour l'IRTF, mais démontrent du potentiel de l'UV-visible pour la prédiction des anthocyanes. Si l'on étudie les résultats de modélisation en fonction du millésime, nous constatons toutefois que celui-ci impacte très fortement la prédiction. Avec un R<sup>2</sup>P de 0,5 et un RPD P de 0,13, ce modèle est inapte à prédire les échantillons de 2018 en fonction de ceux de 2019. Les résultats présentés indiquent donc qu'il est possible de prédire les anthocyanes totales grâce à l'UV-visible, mais qu'il est nécessaire d'avoir une calibration avec une variabilité des millésimes. Ces résultats doivent être approfondis en intégrant d'autres millésimes à la calibration pour confirmer ou infirmer l'intérêt de l'UV-visible dans la prédiction des anthocyanes.

Modèles de prédiction PLS du dosage des anthocyanes moléculaires par IRTF-UV-visible pendant la maturité du raisin

Compte tenu des conclusions qui ont pu être avancées sur la modélisation des anthocyanes moléculaires par UV-visible et IRTF, seuls les résultats des modèles de prédiction pour

les anthocyanes totales seront discutés pour la combinaison des deux analyses spectrales. Les résultats obtenus sont présentés dans le tableau 17.

<b>Validation croisée</b>						
<b>Paramètre</b>	<b>R<sup>2</sup>C</b>	<b>RMSEC</b>	<b>R<sup>2</sup>CV</b>	<b>RMSECV</b>	<b>RPD CV</b>	<b>VL</b>
Total anthocyanes (mg/L)	0,86	28,42	0,81	24,07	2,70	10
<b>Cal/Val Kennard-Stone</b>						
<b>Paramètre</b>	<b>R<sup>2</sup>P</b>	<b>RMSEP</b>	<b>RPD P</b>	<b>VL</b>		
Total anthocyanes (mg/L)	0,93	17,08	3,81	12		
<b>Cal 2019, Val 2018</b>						
<b>Paramètre</b>	<b>R<sup>2</sup>P</b>	<b>RMSEP</b>	<b>RPD P</b>	<b>VL</b>		
Total anthocyanes (mg/L)	0,72	43,10	1,51	2		

Tableau 17 : Modèles de prédiction PLS du dosage des anthocyanes totales par IRTF-UV-visible (R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Les résultats de validation croisée de l'IRTF-UV-visible (R<sup>2</sup>CV : 0,81, RPD CV : 2,70) montrent des valeurs de prédiction supérieures à celles obtenues pour l'UV-visible (R<sup>2</sup>CV : 0,78, RPD CV : 2,15) et inférieures à l'IRTF (R<sup>2</sup>CV : 0,90, RPD CV : 3,16). La combinaison spectrale n'indique donc pas une augmentation de la précision mais une prédiction intermédiaire, où l'addition de l'UV-visible à l'IRTF augmente l'incertitude de la modélisation. Si la validation croisée est un premier indicateur fiable de la robustesse d'un modèle, il ne doit pas remplacer la validation par un jeu de données extérieures à la calibration.

Si l'on s'intéresse aux résultats de la validation avec l'algorithme de Kennard-Stone, nous constatons que la combinaison des deux analyses spectrales augmente la prédiction de l'analyse. Ainsi, nous observons une baisse de l'erreur de prédiction de 2,44 mg/L par rapport à l'IRTF, et de 10,26 mg/L par rapport à l'UV-visible. Rappelons que lors de la division du jeu de données par l'algorithme de Kennard-Stone, celui-ci cherche à présenter un groupe de validation représentatif de la variabilité de la calibration. Ainsi, lorsque le modèle conçu est appliqué à des échantillons similaires à la calibration, la combinaison spectrale apporte un gain de robustesse. Il est alors possible d'envisager que l'UV-visible et l'IRTF puissent être complémentaires dans l'information apportée pour la prédiction des anthocyanes totales. Malheureusement, les résultats de la prédiction par millésime ne permettent pas de valider cette hypothèse. Les valeurs obtenues montrent un RPD P de 1,51, inférieur à l'IRTF seul (RPD P : 1,99).

Nous avons pu constater que l'IRTF pouvait s'affranchir d'une variabilité forte pour la prédiction, notamment celle du millésime, contrairement à l'UV-visible. Si la modélisation est construite avec un

jeu de données assez important pour couvrir la variabilité des échantillons à prédire, la combinaison des deux analyses spectrales pourrait permettre d'augmenter la prédiction. Pour valider l'intérêt de l'IRTF-UV-visible, il est alors nécessaire de travailler sur une calibration composée d'un nombre d'échantillons plus important, couvrant une plus forte variabilité.

***c) Construction de modèles de prédiction PLS pour le suivi du ratio pigments monomériques/pigments polymériques sur les millésimes 2018 et 2019 pendant la maturité du raisin***

Le ratio des pigments monomériques et polymériques a été obtenu par le dosage décrit par Harbertson et al., 2003. Cette mesure a été effectuée sur les 6 parcelles de raisins suivies en 2018, ainsi que sur les 15 parcelles suivies en 2019. Les résultats sont présentés sur la Figure 56. La numérotation des parcelles suivies est identique à celle utilisée pour la méthode Glories.

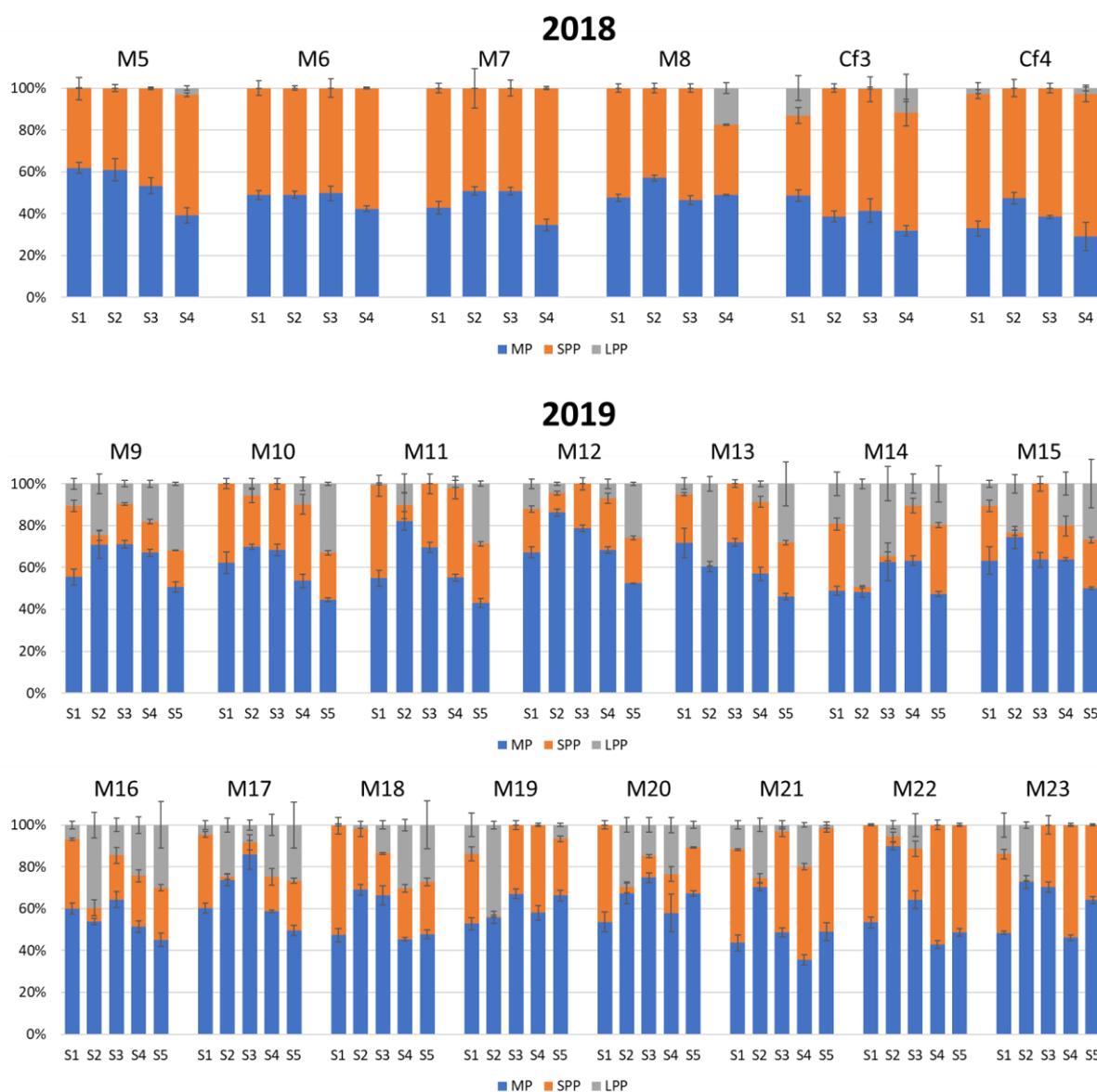


Figure 56 : Suivi du ratio pigments monomériques (MP), petits pigments polymérisés (SPP) et grands pigments polymérisés (LPP) pendant la maturité (S : semaine de prélèvement)

Le ratio des différents pigments est indiqué par parcelle, en fonction de son évolution durant la maturité. Le prélèvement de la semaine 1 représente le premier prélèvement du millésime, et le dernier prélèvement (S4 en 2018 et S5 en 2019) représente l'échantillon lors de la vendange.

Les résultats obtenus sont très hétérogènes, et montrent une grande disparité en fonction des parcelles suivies et du millésime. Nous remarquons des ratios en pigments polymérisés (SPP et LPP) très importants, notamment pour les parcelles Cf3 et Cf4 où les polymères représentent plus de 50 % des pigments mesurés. De plus, nous observons des évolutions très fortes dans la différence de ratio entre les petits pigments polymérisés (SPP) et les grands pigments polymérisés (LPP). Sur la parcelle M14,

nous constatons un ratio en LPP de 20 % en semaine 1, qui évolue jusqu'à 50 % en semaine 2 pour ensuite diminuer à 10 % en semaine 4. Les proportions des pigments polymérisés par rapport aux pigments monomériques peuvent être discutées. Dans la littérature, les résultats de cette analyse sur le raisin montrent des valeurs similaires, alors que des analyses plus fines (utilisation de CLHP couplée à la spectrométrie de masse) indiquent une proportion en pigments polymérisés très inférieure (275,345,346). Les résultats de cette méthode de dosage ne concordent donc pas avec les concentrations réelles mesurées pour les pigments polymérisés du raisin. De plus, nous constatons des écarts-type très importants, qui influencent le résultat final. Cette analyse n'apparaît donc pas assez précise et adaptée pour la mesure des pigments du raisin.

Les modèles par régression PLS pour les paramètres du dosage Adams-Harbertson ont été construits et présentés dans le Tableau 18. Nous avons pu constater lors de l'analyse des anthocyanes moléculaires que les anthocyanes totales étaient prédictibles et que certaines corrélations d'apparence exploitables pouvaient être remises en question. Ainsi, discuter les résultats de ce dosage des pigments nous permet de confirmer qu'il n'est pas adapté au raisin, et d'investiguer la possibilité de sur-corrélation de données qui ne reflètent pas l'évolution des pigments lors de la maturité.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL
<b>UV-visible</b>						
MP (absorbance)	0,69	0,14	0,51	0,18	1,39	9
SPP (absorbance)	0,77	0,18	0,57	0,25	1,48	9
LPP (absorbance)	0,46	0,19	0,30	0,22	1,23	8
PP (absorbance)	0,71	0,20	0,60	0,24	1,58	9
<b>IRTF</b>						
MP (absorbance)	0,73	0,13	0,61	0,16	1,56	11
SPP (absorbance)	0,64	0,22	0,58	0,24	1,54	5
LPP (absorbance)	0,46	0,20	0,31	0,22	1,23	8
PP (absorbance)	0,56	0,25	0,49	0,27	1,41	5
<b>IRTF-UV-visible</b>						
MP (absorbance)	0,66	0,14	0,52	0,17	1,47	11
SPP (absorbance)	0,64	0,22	0,56	0,24	1,54	8
LPP (absorbance)	0,47	0,19	0,26	0,26	1,04	11
PP (absorbance)	0,72	0,20	0,62	0,23	1,65	11

Tableau 18 : Modèles de prédiction PLS du dosage des pigments monomériques (MP), petits pigments polymérisés (SPP) grands pigments polymérisés (LPP) et pigments polymérisés totaux (PP = SPP + LPP) (R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Seuls les résultats de calibration et de validation croisée sont présentés. Les valeurs de prédiction apparaissent faibles, avec un  $R^2CV$  maximum de 0,62 pour les pigments polymérisés totaux par IRTF-UV-visible, et un RPD CV de 1,65. Aucun des modèles n'indique de résultats exploitables, ce qui semble confirmer le caractère inadapté de ce dosage appliqué à la maturité du raisin.

#### ***d) Construction de modèles de prédiction PLS pour le suivi de l'évolution des tanins sur les millésimes 2018 et 2019 pendant la maturité du raisin***

##### ***(1) Evolution des tanins pendant la maturité du raisin***

L'évolution des tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose a été suivie sur la maturité des raisins des millésimes 2018 et 2019. Ces mesures ont été effectuées sur les 6 parcelles de raisins suivies en 2018, ainsi que sur les 15 parcelles suivies en 2019. Les résultats sont présentés en Figure 57. La numérotation des parcelles suivies est identique à celle utilisée pour la méthode Glories.

Nous constatons une évolution différente suivant la méthode d'analyse utilisée. Si ces deux dosages reposent sur la précipitation des tanins, la différence entre protéines et polysaccharides permet de différencier deux types d'évolution. Cependant, certains comportements sont comparables pour les deux analyses. Nous observons ainsi sur le millésime 2018 une diminution globale de la teneur en tanin de la semaine 1 à la semaine 3, suivie d'une augmentation jusqu'à la vendange. La même évolution est constatable sur le millésime 2019. Nous remarquons toutefois des divergences entre les deux études. Pour le millésime 2018, l'augmentation de la concentration des tanins précipitables à la méthylcellulose (MCP) est observable pour pratiquement toutes les parcelles suivies, à l'exception de la parcelle Cf4. Pour les tanins précipitables à l'albumine de sérum bovin (BSA) cette observation n'est applicable que pour les parcelles M8 et Cf3. Pour le millésime 2019, l'évolution de MCP est similaire pour pratiquement toutes les parcelles, avec une diminution des concentrations progressive entre la semaine 1 et la semaine 3, suivie d'une augmentation de la semaine 3 à la semaine 5. Seule la parcelle M23 apparaît ne pas suivre cette tendance. Pour BSA, nous constatons une évolution plus hétérogène entre les parcelles, avec les parcelles M15, M16, M18, M20 et M21 qui évoluent différemment de la tendance globale.

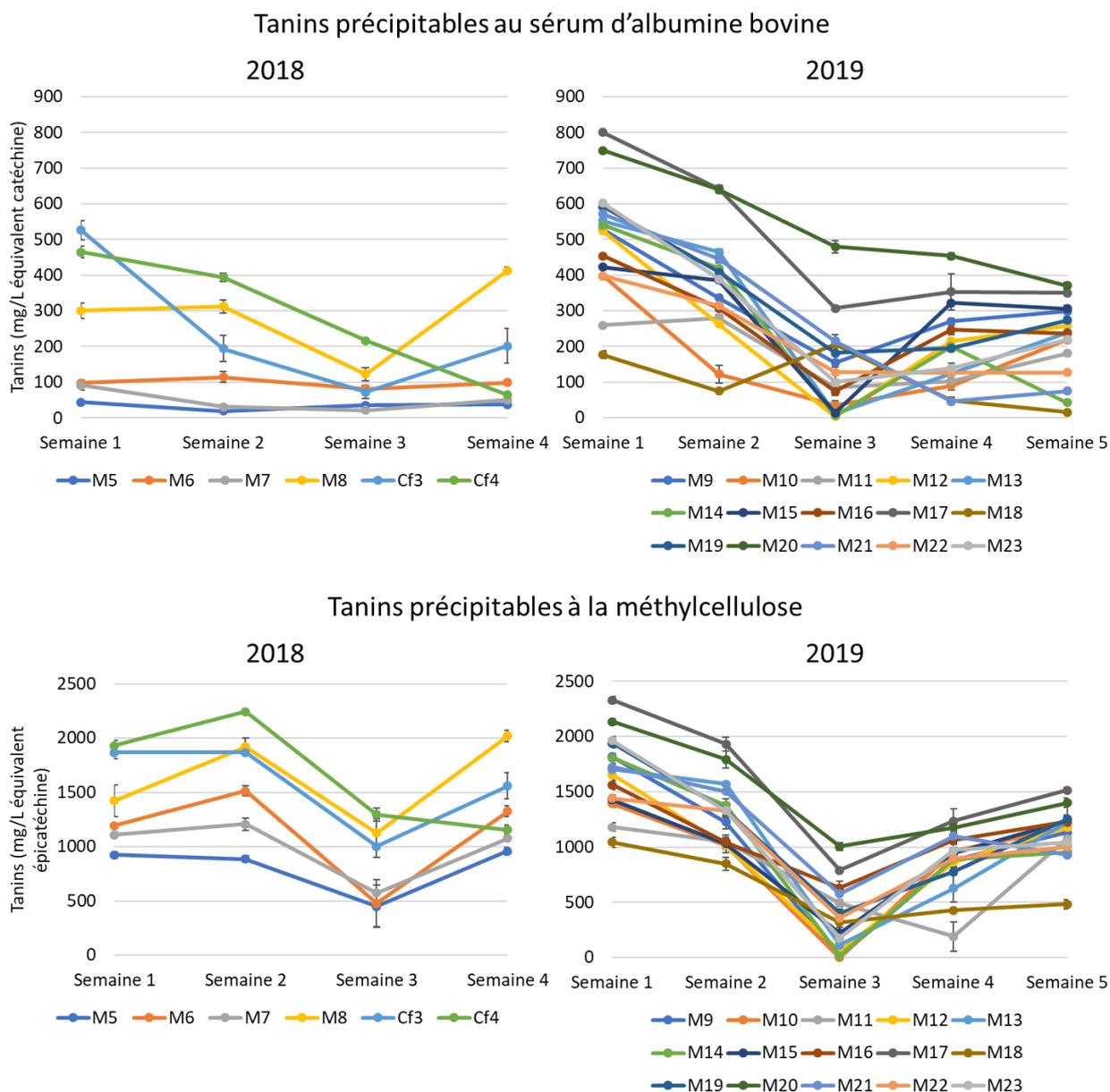


Figure 57 : Evolution de la concentration en tanins sur les millésimes 2018 et 2019 par dosage à l'albumine de sérum bovin et à la méthylcellulose (M : Merlot, Cf : Cabernet franc)

Après la véraison, l'accumulation des tanins dans le pépin et la pellicule n'évolue que très peu (97,347,348). Toutefois, le niveau de maturité va influencer l'extractibilité de ces composés (349). De plus, le volume de la baie est aussi sujet à des variations influencées par les conditions climatiques et la maturité, ce qui conduit à plus ou moins diluer les concentrations en tanins. Dès lors, il est intéressant pour le vinificateur de pouvoir suivre l'évolution de la concentration en tanins extractibles. L'utilisation de deux dosages différents permet de différencier les tanins réactifs aux protéines (BSA),

fortement corrélé avec l'astringence, des tanins réactifs à la méthylcellulose (MCP) qui sont moins spécifiques et permettent d'obtenir une concentration des tanins totaux (275,325,327,337).

(2) *Conception et validation de modèles de prédiction PLS pour le dosage des tanins*

En utilisant la régression PLS, des modèles de prédiction sont construits pour l'UV-visible, l'IRTF, et l'IRTF-UV-visible. Le Tableau 19 présente les résultats de prédictions obtenues en utilisant tout le jeu de données.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL
<b>UV-visible</b>						
BSA (mg/L)	0,86	68,84	0,82	80,18	2,35	6
MCP (mg/L)	0,87	196,03	0,82	231,62	2,33	9
<b>IRTF</b>						
BSA (mg/L)	0,91	57,29	0,87	68,55	2,75	9
MCP (mg/L)	0,82	230,67	0,78	253,90	2,13	9
<b>IRTF-UV-visible</b>						
BSA (mg/L)	0,83	78,36	0,78	87,69	2,15	8
MCP (mg/L)	0,67	310,20	0,58	353,04	1,53	8

Tableau 19 : Modèles de prédiction PLS pour les tanins précipitables à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP) par UV-visible, IRTF et IRTF-UV-visible. Utilisation du jeu de données complet pour la calibration (R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Les données obtenues pour l'UV-visible et l'IRTF sont très proches. La validation croisée apporte des résultats robustes et prédictifs, et montre la faculté des deux analyses spectrales à prédire les concentrations en tanins extractibles lors de la maturité. Pour le dosage à l'albumine de sérum bovin (BSA), l'IRTF apparaît plus performant, avec un R<sup>2</sup>CV de 0,87 et un RPD CV de 2,75 comparé à l'UV-visible avec un R<sup>2</sup>CV de 0,82 et un RPD CV de 2,35. Inversement, pour le dosage à la méthylcellulose (MCP), c'est l'UV-visible qui est plus robuste, avec R<sup>2</sup>CV de 0,82 et un RPD CV de 2,33 comparé à l'IRTF avec un R<sup>2</sup>CV de 0,78 et un RPD CV de 2,13. Cette dernière différence reste relativement faible, les deux analyses spectrales sont donc très proches dans leur prédictivité.

Les résultats de l'IRTF-UV-visible sont plus faibles que l'UV-visible et l'IRTF seul pour les deux paramètres des tanins. La combinaison ne permet donc pas de renforcer la prédiction, mais apporte au contraire plus d'incertitude au modèle. Afin de confirmer les résultats observés, les paramètres de

prédiction pour la validation utilisant l'algorithme de Kennard-Stone et la validation par millésime sont présentés dans le Tableau 20.

Paramètre	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL
Cal/Val Kennard-Stone					Cal 2019, Val 2018			
<b>UV-visible</b>								
BSA (mg/L)	0,84	84,26	2,24	7	0,20	288,03	0,65	6
MCP (mg/L)	0,84	218,48	2,47	8	0,09	611,45	0,88	8
<b>IRTF</b>								
BSA (mg/L)	0,87	82,25	2,29	8	0,73	179,47	1,05	9
MCP (mg/L)	0,79	239,90	2,25	10	0,21	649,33	0,83	9
<b>IRTF-UV-visible</b>								
BSA (mg/L)	0,87	77,44	2,43	8	0,36	231,28	0,81	6
MCP (mg/L)	0,64	276,42	1,95	6	0,28	440,65	1,22	8

Tableau 20 : Résultats de prédiction PLS avec validation par échantillons externes pour les tanins précipitables à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP) par UV-visible, IRTF et IRTF-UV-visible (R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Lorsque le jeu de données utilisé pour la calibration et la validation présente la même variabilité (division des données par l'algorithme de Kennard-Stone), les résultats observés montrent une prédictivité similaire aux résultats obtenus par validation croisée. Pour l'UV-visible et l'IRTF, le RPD P indique des valeurs supérieures à 2 pour les deux dosages des tanins, avec respectivement 2,24 et 2,29 pour BSA, et 2,47 et 2,25 pour MCP. Ainsi, les deux analyses spectrales permettent la prédiction de ces paramètres avec une erreur de validation jugée acceptable. Cependant, de tels modèles ne doivent être utilisés que pour obtenir une valeur indicative de la teneur en tanins, et non comme une réelle mesure de la concentration. Si l'on se réfère à la littérature, ces modèles nécessiteraient un RPD P supérieur à 3 pour être considérés comme fiables sur la valeur de la concentration (10). Néanmoins, cela démontre le potentiel de l'UV-visible et de l'IRTF pour la prédiction des tanins lors de la maturité du raisin. Concernant l'IRTF-UV-visible, la valeur de RPD P de BSA (2,43) s'avère légèrement supérieure à celle présentée précédemment, alors que la valeur pour MCP (1,95) est quant à elle inférieure. L'augmentation reste très proche des valeurs pour l'IRTF et l'UV-visible, et en contradiction avec la validation croisée. Il est difficile d'attribuer cette hausse de robustesse à un effet de combinaison, et dans l'ensemble, le gain de prédiction est négligeable par rapport aux instruments nécessaires pour effectuer l'analyse. Une mesure UV-visible ou IRTF seule est alors préférable.

Si l'on s'intéresse à la validation par millésime, nous constatons une très forte diminution de la prédiction. Pour l'UV-visible, les RPD P sont de 0,65 pour BSA et 0,88 pour MCP. Pour l'IRTF, les RPD P sont de 1,05 pour BSA et 0,83 pour MCP. Ces résultats finissent de valider le fort effet millésime observable sur les spectres, déjà observé pour les anthocyanes. Compte tenu de l'ACP des spectres UV-visible présenté en Figure 45, cet effet s'explique par le manque de variabilité exprimé sur le millésime de calibration, qui ne permet pas de prédire le millésime de validation. Pour l'ACP de l'IRTF (Figure 46), ce manque de variabilité n'est pas observé et ne permet pas d'expliquer l'impact du millésime. Cependant, l'information spectrale contenue dans l'IRTF est bien plus riche, et contient les bandes d'absorption de la majorité des composés organiques présente dans le vin. Nous pouvons alors émettre l'hypothèse que les variations de composés tels que les sucres ou les acides qui sont sujet à des différences d'évolution en fonction des années influencent la prédiction.

### **3. Conclusions et perspectives de la conception de modèles de prédiction PLS pour l'analyse des anthocyanes et tanins du raisin pendant la maturité du raisin**

L'objectif principal de cette première partie était d'étudier la possibilité d'utiliser l'UV-visible et l'IRTF couplé à l'analyse de données multivariée pour prédire l'évolution des anthocyanes et des tanins pendant la maturité du raisin. Pour y parvenir, nous avons utilisé la méthode d'extraction à pH 3,2 décrite par Glories pour obtenir des échantillons représentatifs de l'extractibilité de ces composés. Nous avons enregistré les spectres de ces échantillons et en parallèle, nous avons utilisé des analyses de références décrites dans la littérature pour obtenir différents paramètres permettant d'expliquer l'évolution des anthocyanes et des tanins.

Les résultats obtenus nous ont permis de mettre en évidence le potentiel des deux analyses spectrales pour prédire les anthocyanes totales et les tanins. La comparaison de ces deux analyses indique des prédictions plus robustes obtenues avec l'IRTF, en particulier pour les anthocyanes totales. De plus, la combinaison IRTF-UV-visible a démontré pour ce paramètre des prédictions supérieures, et pourrait mettre en évidence un effet complémentaire des deux spectres. Cependant, cette hypothèse n'a pas pu être vérifiée avec les autres paramètres étudiés.

Nous avons également pu démontrer les limites de l'analyse multivariée appliquée à une matrice aussi complexe que celle du raisin. En effet, si les résultats obtenus pour chaque anthocyane moléculaire suggèrent la possibilité de prédire leur concentration, nous avons pu prouver que les corrélations entre

paramètres pouvaient influencer fortement un modèle et conduire à une mauvaise interprétation des résultats. Une étude approfondie du jeu de données étudié est primordiale, afin d'obtenir des modèles performants construits sur l'empreinte spectrale du composé recherché.

L'UV-visible et l'IRTF apparaissent tout à fait adaptés pour suivre l'évolution des anthocyanes et des tanins du raisin pendant la maturité. Cependant, les données recueillies lors de cette étude ne sont pas suffisantes pour construire des modèles de prédiction efficaces et doivent être complétées. Nous avons pu constater que la variabilité induite par le millésime influence fortement la prédiction. D'autres millésimes, ainsi que d'autres cépages doivent être ajoutés pour obtenir un modèle de prédiction capable de s'adapter à la variabilité importante observée dans le monde viticole. De plus, il est nécessaire de travailler sur une méthode d'extraction des anthocyanes et des tanins efficace et rapide qui reste représentative de l'échantillon étudié, afin de réduire le temps d'analyse. Si un milieu aqueux est suffisant pour extraire les anthocyanes, il est nécessaire d'augmenter la concentration en éthanol pour extraire les tanins. Il est aussi envisageable d'augmenter la température pour accélérer l'extraction et réduire le temps d'analyse.

## **B. Prédiction des concentrations en anthocyanes et tanins du moût et du vin rouge pendant la vinification**

Les concentrations en anthocyanes et tanins du moût et du vin ont été suivies lors de la fermentation alcoolique pour trois millésimes (2017, 2018 et 2019) de l'encuvage jusqu'à l'écoulage. L'ensemble des 21 cuves de vinification suivies sont numérotées de M1 à M15 pour le Merlot, et de Cf1 à Cf6 pour le Cabernet franc. Tout comme pour la conception des modèles pour le raisin, nous avons dans un premier temps étudié la variabilité du jeu de données ainsi que l'évolution dans le temps des paramètres étudiés.

### **1. Analyse du jeu de données.**

#### ***a) Evaluation de la variabilité des analyses de références***

Le jeu de données est représenté par 255 échantillons de moûts et vins rouges, répartis sur trois millésimes (105 en 2017, 60 en 2018 et 90 en 2019). Pour tous les échantillons, les analyses suivantes ont été effectuées :

- L'analyse de la composition en flavan-3-ols par CLHP-UV-visible après réaction de phloroglucinolyse. Cela permet d'obtenir la concentration en unités (monomère après dépolymérisation) catéchine, épicatechine, épigallocatechine et épicatechine-gallate, ainsi que la concentration en unités flavan-3-ols terminales et totales.
- Le dosage des anthocyanes moléculaires par CLHP-UV-visible.
- L'analyse de l'évolution des anthocyanes, avec le dosage d'Adams-Harbertson. Nous avons également réalisé le dosage des anthocyanes libres par décoloration au bisulfite de sodium.
- Le dosage des tanins par précipitation à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP).

Les variabilités de ces analyses sont présentées dans le Tableau 21. Aucun échantillon n'a été écarté.

Paramètre	Moyenne	Min	Max	Ecart-type	CV	Ecart-type Ref
<b>Composition en flavan-3-ols</b>						
Flavan-3-ol (mmol/L)	12,71	1,04	32,34	6,78	53 %	0,72
Flavan-3-ol terminal (mmol/L)	4,16	0,43	11,52	2,14	51 %	0,54
Catéchine (mmol/L)	2,74	0,29	6,96	1,32	48 %	0,20
Epicatéchine (mmol/L)	6,34	0,30	18,28	3,74	59 %	0,27
Epigallocatechine (mmol/L)	3,10	0,34	8,82	1,74	56 %	0,48
Epicatéchine-gallate (mmol/L)	0,54	0,06	2,20	0,36	67 %	0,03
<b>Anthocyanes moléculaires</b>						
Dp-3-O-glc (mg/L)	22,79	0,00	63,26	12,95	57 %	0,63
Cy-3-O-glc (mg/L)	4,90	0,06	13,54	2,87	59 %	0,43
Pt-3-O-glc (mg/L)	24,97	0,00	74,71	12,31	49 %	2,41
Pn-3-O-glc (mg/L)	23,68	0,53	49,51	8,58	36 %	2,79
Mlv-3-O-glc (mg/L)	137,00	1,63	267,86	49,06	36 %	4,01
Dp-3-O-acglc (mg/L)	6,81	0,00	15,03	3,39	50 %	0,28
Cy-3-O-acglc (mg/L)	2,80	0,06	6,64	1,17	42 %	0,43
Pt-3-O-acglc (mg/L)	6,25	0,00	11,97	2,51	40 %	0,40
Pn-3-O-acglc (mg/L)	8,10	0,17	16,55	2,96	37 %	0,44
Mlv-3-O-acglc (mg/L)	42,99	0,53	99,47	16,07	37 %	1,60
Pn-3-O-cmglc (mg/L)	4,69	0,04	9,62	2,24	48 %	0,23
Mlv-3-O-cmglc (mg/L)	18,65	0,07	40,02	8,15	44 %	0,98
Total glucoside (mg/L)	213,35	2,49	411,77	77,58	36 %	5,85
Total acétyl (mg/L)	66,95	0,88	137,89	23,63	35 %	2,24
Total coumaroyl (mg/L)	23,34	0,11	48,03	10,07	43 %	1,17
Total anthocyanes (mg/L)	303,64	3,49	556,17	107,88	36 %	8,80
<b>Evolution des anthocyanes</b>						
MP (Absorbance)	1,57	0,13	3,02	0,65	41 %	0,04
SPP (Absorbance)	0,55	0,00	1,28	0,22	41 %	0,02
LPP (Absorbance)	0,22	0,00	0,90	0,17	76 %	0,03
PP (Absorbance)	0,76	0,06	1,79	0,34	45 %	0,02
Anthocyanes libres (mg/L)	448,15	23,63	721,20	134,97	30 %	6,84
<b>Tanins</b>						
BSA (mg/L)	365,74	2,05	883,34	230,57	63 %	8,47
MCP (mg/L)	1069,42	0,00	2973,06	649,80	61 %	59,35

Tableau 21 : Paramètres des analyses de références obtenus sur les échantillons de moûts et vins en fermentation (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, CV : coefficient de variation)

Pour la concentration en flavan-3-ols, nous constatons une proportion majoritaire de l'epicatechine, qui compose environ 50 % des unités totales. Les concentrations en épigallocatechine et catéchine composent respectivement 24 % et 22 % des unités totales. L'epicatechine-gallate montre la

concentration la plus faible, avec 4 % de la moyenne des unités totales. Ces résultats sont en accord avec les proportions décrites dans la littérature (109). Si l'on s'intéresse à la variabilité de ces composés, nous remarquons des valeurs différentes en fonction de l'unité étudiée, avec une évolution des concentrations différente. L'épicatéchine-gallate possède le coefficient de variation le plus élevé (67 %), ce qui pourrait être dû au fait que cette unité est un marqueur du pépin de raisin. Ainsi, l'extraction plus ou moins importante des tanins du pépin peut impacter sa variabilité plus fortement que pour les autres unités flavan-3-ols. Le coefficient de variation le plus faible est obtenu pour les concentrations en catéchine (48 %), on note néanmoins une variabilité importante du jeu de données.

La variabilité des anthocyanes moléculaires pendant la vinification apparaît légèrement plus importante que pour le suivi du raisin pendant la maturité. Pour la somme des anthocyanes moléculaires, le coefficient de variation passe de 34 % pour le raisin à 36 % pour les vins en fermentation. Si cette différence est faible, nous remarquons cependant des écarts plus importants pour certaines anthocyanes en particulier. Ainsi, la delphinidine-3-*O*-glucoside et la cyanidine-3-*O*-glucoside montrent des coefficients de variabilité respectifs de 57 % et 59 % pour les vins en fermentation, alors qu'ils n'étaient que de 48 % et 51 % pour le raisin. Il n'est cependant pas encore possible d'affirmer que l'évolution de chaque anthocyane moléculaire est indépendante des autres, cet aspect devra être approfondi en étudiant l'évolution des concentrations dans le temps. Si l'on s'intéresse au minimum et maximum de concentration pour chacune des anthocyanes ainsi que pour les anthocyanes totales, nous constatons que le minimum obtenu pour les vins en fermentation est inférieur à celui du raisin, tandis que le maximum est supérieur. Cet écart plus important indique une évolution très marquée pendant la fermentation, qui peut s'expliquer par des concentrations faibles en début de macération pré-fermentaire, et fortes au maximum de l'extraction. Cette évolution dans le temps sera détaillée ultérieurement.

Concernant les paramètres d'évolution des pigments monomériques et polymérisés, nous observons la même variance pour les pigments monomériques (MP) et petit pigments polymérisés (SPP) de 41 %. Ce coefficient de variation apparaît supérieur pour les grands pigments polymérisés (LPP), avec 76 %. Contrairement aux paramètres du raisin, pour les vins en fermentation nous constatons une proportion en MP bien supérieure en SPP et LPP, qui semble mieux refléter la réalité. Cette hypothèse devra être vérifiée avec l'étude de l'évolution pendant la fermentation. Les concentrations en anthocyanes libres montrent quant à elles le coefficient de variation le plus faible tous paramètres confondus (30 %). Cette valeur reste cependant assez élevée pour considérer une bonne répartition des concentrations des échantillons, représentative de la variabilité de la fermentation. Nous

remarquons notamment un écart important entre le minimum de 23,63 mg/L et le maximum de 721,20 mg/L.

L'étude des tanins pendant la fermentation par précipitation à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP) indique des coefficients de variations très similaires, avec respectivement 63 % et 61 %, contrairement à l'étude des tanins pendant la maturité du raisin. Il est possible que l'évolution de ces deux paramètres soit mieux corrélée que pendant l'étude des raisins, ce qui expliquerait cette similitude. Cette information devra être vérifiée par l'analyse en composante principale (ACP).

Tous les paramètres étudiés démontrent une variabilité importante, ce qui est nécessaire pour la conception de modèles de prédiction efficaces. Nous devons désormais étudier les possibles corrélations entre paramètres et individus grâce à l'ACP. Cette analyse permettra de mettre en évidence les similitudes entre les analyses de références ainsi que les groupes d'échantillons.

### ***b) Analyse en composantes principales (ACP) des analyses de références et des données spectrales***

#### ***(1) ACP des variables des analyses de références***

Deux ACP des variables des analyses de références sont présentées sur la Figure 58. En (1) l'ACP est construite avec tous les paramètres étudiés, à l'exception des anthocyanes moléculaires pour éviter que ceux-ci n'influencent trop fortement l'analyse. Cette ACP permet d'étudier les possibles corrélations entre paramètres. En (2) l'ACP est construite avec seulement les anthocyanes moléculaires et la somme des anthocyanes, ainsi que la somme des anthocyanes-3-*O*-glucoside, la somme des anthocyanes-3-*O*-acétylglucoside et la somme des anthocyanes-3-*O*-coumaroylglucoside. Cette ACP permet d'évaluer la répartition des concentrations en anthocyanes moléculaires entre elles. Nous avons pu constater avec l'analyse de ces concentrations pendant la maturité du raisin que de fortes corrélations pouvaient être observées, et impacter la conception des modèles. Ainsi, il est nécessaire d'évaluer si les concentrations en anthocyanes sont indépendantes les unes des autres ou fortement corrélées, pour mettre en évidence ou non la possibilité de corrélation indirecte.

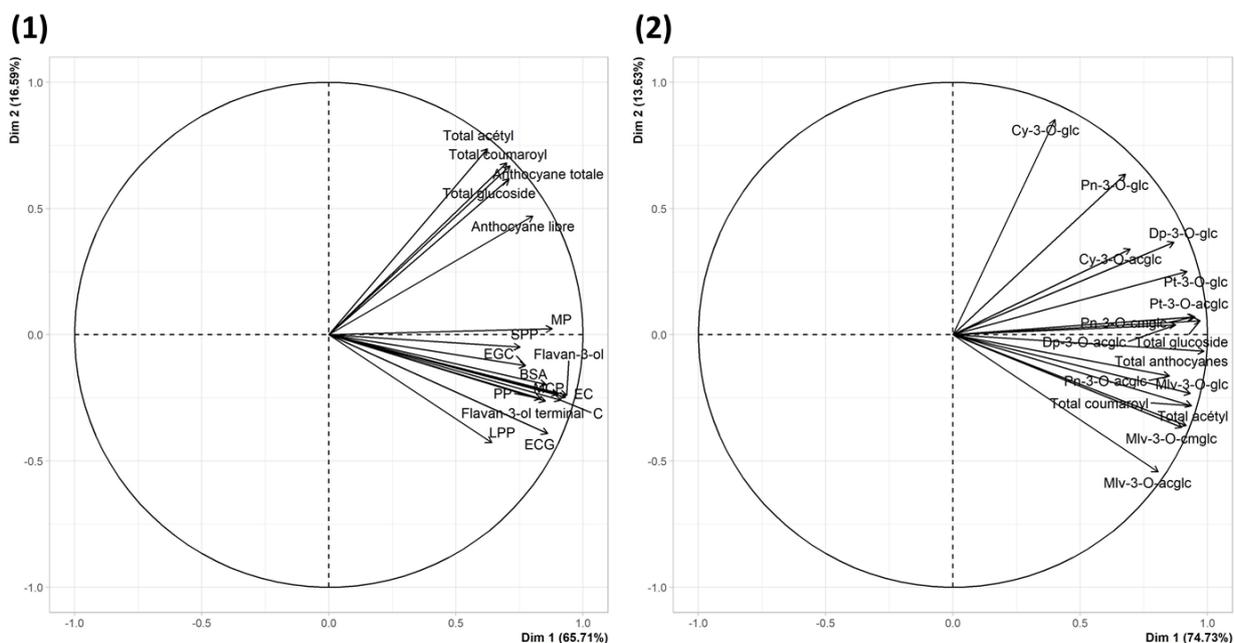


Figure 58 : ACP des variables des analyses de références des moûts et vins en fermentation. (1) ACP de tous les paramètres à l'exception des concentrations en anthocyanes moléculaires individuelles, (2) ACP des concentrations anthocyanes moléculaires individuelles et des sommes des anthocyanes (totale, glucoside, acétyl et coumaroyl) (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, C : catéchine, EC : épicatechine, EGC : épigallocatechine, ECG : épicatechine-gallate)

Sur l'ACP (1), la dimension 1 explique 65,71 % de la variabilité et la dimension 2 16,59 %, pour un total de 82,30 %. Nous constatons également que toutes les variables sont proches du cercle des corrélations, et sont donc correctement expliquées par l'analyse. Les variables se divisent en deux groupes : le premier est constitué de la plupart des analyses de références pour doser les anthocyanes : les anthocyanes libres, la somme des anthocyanes moléculaires, la somme des anthocyanes-3-O-glucoside, la somme des anthocyanes-3-O-acétylglucoside et la somme des anthocyanes-3-O-coumaroylglucoside. Le deuxième groupe est composé des variables qui correspondent aux tanins et aux anthocyanes polymérisées, soit les concentrations en tanins précipitables à la méthylcellulose (MCP) et l'albumine de sérum bovin (BSA), les concentrations en unités catéchine, épicatechine, épigallocatechine, épicatechine-gallate, terminale et totale, ainsi que les teneurs en petits et grands pigments polymérisés. Nous constatons que la teneur en pigments monomériques est la variable la mieux corrélée avec la dimension 1, et qu'elle se détache légèrement du groupe de variables des tanins pour se rapprocher du groupe de variables des anthocyanes. Ces résultats indiquent donc une discrimination entre l'évolution des anthocyanes et des tanins lors de la fermentation.

Sur l'ACP (2), la dimension 1 explique 74,73 % de la variabilité et la dimension 2 13,63 %, pour un total de 88,36 %. Toutes les variables des concentrations en anthocyanes moléculaires sont positivement corrélées avec la dimension 1, avec des divergences plus importantes sur la dimension 2. En comparant cette ACP avec celle des concentrations en anthocyanes moléculaires lors de la maturité du raisin (Figure 43), nous remarquons de très fortes similitudes qui laissent à penser que la distribution des anthocyanes suit le même schéma. Il sera donc nécessaire de s'assurer que les modèles construits pour doser les concentrations en anthocyanes moléculaires sont indépendants les uns des autres, afin d'éviter, comme pour la maturité du raisin, une corrélation indirecte avec la somme des anthocyanes moléculaires.

*(2) ACP des individus en fonction des analyses de références et des spectres UV-visible et IRTF*

L'ACP des individus en fonction des analyses de références est représentée sur la Figure 59. Tous les paramètres ont été utilisés à l'exception des concentrations en anthocyanes moléculaires individuelles. Les données sont identiques à celles de l'ACP (1) de la Figure 58, les ACP des variables et des individus sont donc comparables entre elles. Chaque point représente un échantillon, et les données utilisées sont la moyenne du triplicat d'analyse.

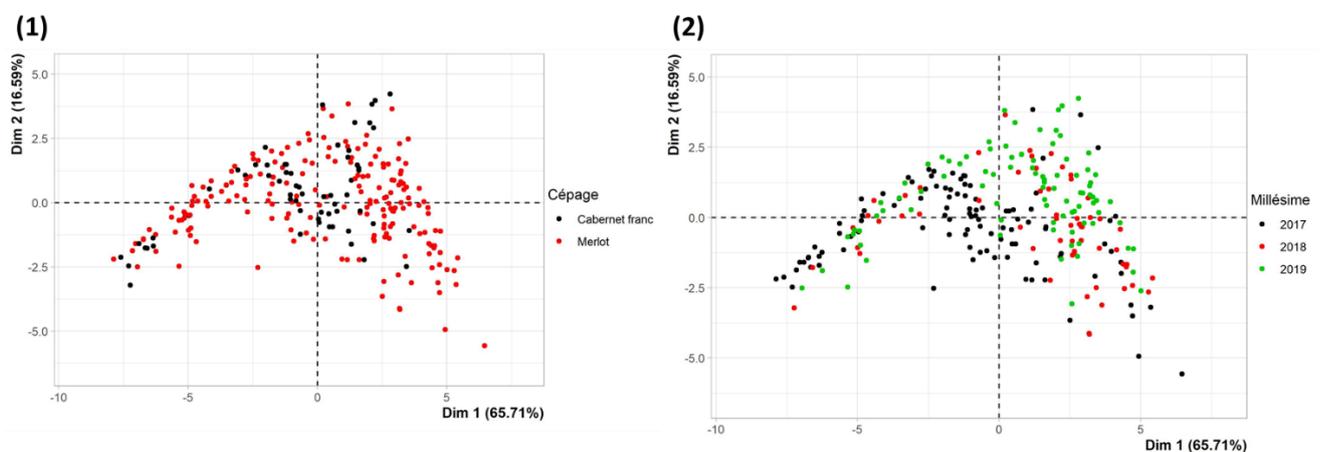


Figure 59 : ACP des individus des analyses de références : (1) représentation en fonction du cépage (2) représentation en fonction du millésime

Nous constatons une répartition des échantillons du nuage de points homogène, avec très peu d'individus qui se détachent de la masse et aucun regroupement apparent. Cette distribution apparaît idéale pour concevoir des modèles de prédiction, car les concentrations mesurées couvrent toute la diversité de la fermentation alcoolique. Si l'on s'intéresse à la distribution par cépage de l'ACP (1),

nous constatons que les échantillons de Cabernet franc et de Merlot sont répartis de manière homogène. Pour les concentrations en anthocyanes et tanins des analyses de références, les cépages étudiés ne semblent pas discriminants. A contrario, sur l'ACP (2) par millésime, nous distinguons une répartition plus hétérogène, notamment avec les échantillons du millésime 2017 qui se détachent du reste des données. Cette discrimination est faite en suivant l'axe de corrélation des différents paramètres relatifs aux anthocyanes (Figure 58) et indique donc une divergence observable dans les concentrations en anthocyanes en fonction du millésime. Tout comme pour l'analyse des raisins pendant la maturité, l'impact du millésime est important, et il devra donc être pris en compte lors de la conception des modèles afin de s'assurer que ceux-ci garderont leur performance d'une année à l'autre.

L'ACP des individus en fonction des spectres UV-visible est représentée sur la Figure 60. La dimension 1 représente 75,56 % de la variabilité, et la dimension 2 8,20 %, pour un total de 83,76 %.

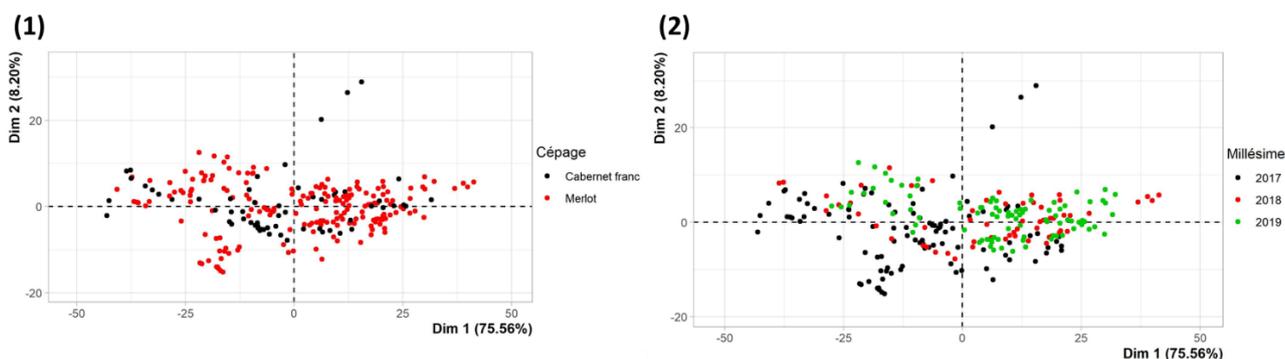


Figure 60 : ACP des individus des spectres UV-visible : (1) représentation en fonction du cépage (2) représentation en fonction du millésime

Le nuage de points montre une distribution homogène sur la dimension 1. A l'inverse, sur la dimension 2, plusieurs échantillons du millésime 2017 se détachent légèrement du groupe principal d'individus. Comme ceux-ci apparaissent excentrés seulement sur la dimension 2, qui n'exprime que 8,20 % de la variabilité, il n'est pas nécessaire de les considérer comme déviants pour le moment. Ce résultat peut être attribué à des spectres atypiques par rapport au jeu de données complet, mais aussi à une erreur de lecture ou de manipulation. Par ailleurs, il s'agissait du millésime le plus discriminé sur l'ACP des analyses de références, et cela apparaît donc en accord avec l'ensemble des données. Si lors de la conception des modèles nous constatons des prédictions plus faibles pour le millésime 2017, il pourra être envisagé de les exclure.

Si l'on compare cette ACP pour les échantillons en fermentation à l'ACP des spectres UV-visible des raisins pendant la maturité (Figure 45), nous remarquons que la répartition est bien plus homogène, et semble mieux adaptée à la conception de modèles de prédiction. De plus, en s'intéressant à la répartition par millésimes et par cépages, nous n'observons pas de distinction. Le paramètre du millésime, notamment, qui influait fortement la prédiction pendant la maturité, semble moins impactant lors de la fermentation.

L'ACP des individus en fonction des spectres IRTF est représentée sur la Figure 61. La dimension 1 explique 88,18 % de la variabilité et la dimension 2 10,44, pour un total de 98,62 %. Les deux premières dimensions expriment très fortement la variabilité des spectres IRTF.

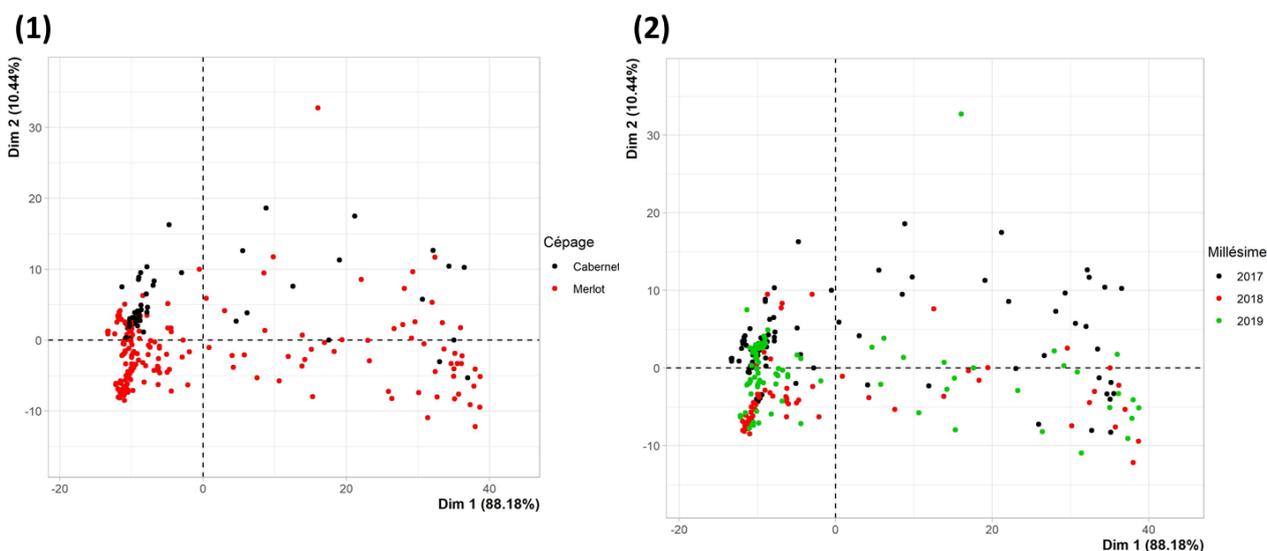


Figure 61 : ACP des individus des spectres IRTF : (1) représentation en fonction du cépage (2) représentation en fonction du millésime

Concernant la répartition des échantillons, celle-ci est plus hétérogène que pour les analyses de références et les spectres UV-visible. Sur la dimension 1, nous remarquons un regroupement majoritaire d'échantillons, avec une répartition plus dispersée pour la minorité des individus restants. Si l'on s'intéresse à la représentation par cépage et par millésime, nous constatons une légère discrimination sur la dimension 2 mais celle-ci ne permet pas d'expliquer la dispersion du nuage de points sur la dimension 1. Il faut rappeler que la plupart des composés organiques de la matrice du moût et du vin absorbent dans l'infrarouge. Lors de la fermentation, cette matrice évolue fortement, notamment avec la diminution de la concentration en sucre et l'augmentation de la concentration en éthanol qui peuvent impacter l'empreinte spectrale. Afin d'évaluer l'impact de ces changements sur

la variabilité des spectres IRTF, l'ACP des individus en fonction des jours de fermentation est représentée sur la Figure 62.

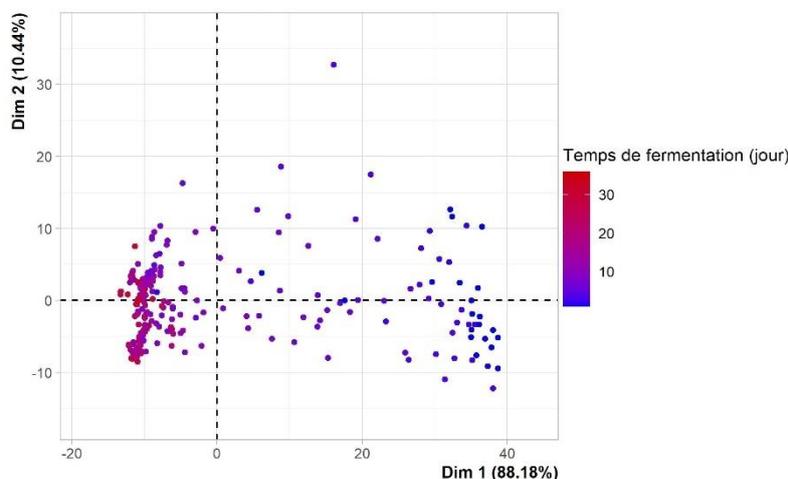


Figure 62 : ACP des individus des spectres IRTF : représentation en fonction du temps de fermentation

Nous observons l'impact important des changements pendant la fermentation alcoolique sur la variabilité des spectres IRTF. Les échantillons sont répartis sur la dimension 1 en fonction du nombre de jours de vinification. Ainsi, la différence de cépage et de millésime n'a que peu d'impact sur la variabilité, celle-ci étant due principalement à l'évolution de la fermentation.

Dans l'ensemble, le jeu d'échantillons étudié apparaît idéal pour concevoir des modèles de prédiction, avec une répartition qui permet de couvrir tout le spectre des concentrations étudiées et très peu d'échantillons déviants. Nous constatons également que l'impact du millésime et du cépage est faible, ce qui peut être un bon indicateur pour concevoir des modèles qui s'adapteront facilement à des échantillons extérieurs à la calibration. Nous notons cependant que les spectres IRTF apparaissent très impactés par l'évolution de la matrice du moût et du vin pendant la fermentation, ce qui peut influencer la précision des modèles créés.

### ***c) Analyse discriminante PLS pour la différenciation des millésimes et des cépages étudiés***

Tout comme pour l'étude des raisins pendant la maturité, nous avons utilisé l'analyse discriminante PLS (PLSDA) afin d'évaluer la possibilité de distinguer des groupes d'échantillons par millésime ou par cépage. Dans le Tableau 22 sont répertoriés les résultats obtenus pour la discrimination de ces groupes, avec le coefficient de calibration et de validation croisée ( $R^2C$  et  $R^2CV$ ), ainsi que le nombre de variables latentes (VL) utilisées. Comme nous disposions de trois

millésimes, nous avons pu essayer de discriminer chaque millésime par rapport aux deux autres, ce qui explique pourquoi nous présentons trois modèles PLS-DA différents.

<b>Groupe discriminé</b>	<b>R<sup>2</sup>C</b>	<b>R<sup>2</sup>CV</b>	<b>VL</b>
<b>Analyses de références</b>			
Millésime 2017	0,95	0,94	8
Millésime 2018	0,97	0,95	8
Millésime 2019	0,94	0,91	8
Cépage	0,83	0,81	4
<b>Spectres UV-visible</b>			
Millésime 2017	0,88	0,87	7
Millésime 2018	0,83	0,79	7
Millésime 2019	0,89	0,85	7
Cépage	0,54	0,58	3
<b>Spectres IRTF</b>			
Millésime 2017	0,83	0,79	4
Millésime 2018	0,72	0,72	4
Millésime 2019	0,84	0,81	4
Cépage	0,78	0,75	3

Tableau 22 : Analyses PLS discriminantes pour la différenciation des millésimes et des cépages du jeu de données (R<sup>2</sup>C : coefficient de régression de la calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, VL : nombre de variables latentes)

Nous constatons que les discriminations sont bien moins performantes que pour l'analyse des raisins pendant la maturité. La meilleure séparation est obtenue grâce aux analyses de références, ce qui est corrélé avec les observations qui avaient été faites sur les ACP des individus. Les trois millésimes montrent des résultats relativement proches, avec un R<sup>2</sup>CV de 0,94 pour 2017, de 0,95 pour 2018 et de 0,91 pour 2019. Les analyses de références apparaissent donc discriminantes pour le millésime, avec cependant une incertitude qui ne permet pas une séparation parfaite. Pour les cépages, la différenciation est moins importante, avec un R<sup>2</sup>CV de 0,81. Ainsi, pour les données recueillies avec les analyses de références, nous constatons que le millésime impacte les concentrations en tanins et en anthocyanes. A l'inverse, les cépages utilisés ne semblent pas influencer fortement sur la répartition de ces concentrations. Dans l'ensemble, ces résultats démontrent d'un effet millésime proche de celui obtenu sur le suivi de la maturité (R<sup>2</sup>CV de 0,91), ainsi que d'un effet cépage plus faible.

Pour les spectres UV-visible et IRTF, la robustesse de la discrimination est inférieure à celle des analyses de références. La meilleure séparation pour les spectres UV-visible est obtenue pour le millésime 2017 (R<sup>2</sup>CV de 0,87) et pour les spectres IRTF pour le millésime 2019 (R<sup>2</sup>CV de 0,81). Tout comme pour les analyses de références, les R<sup>2</sup>CV des millésimes sont proches, il n'y a donc pas

de distinction forte d'une année en particulier. Concernant la distinction des cépages, les résultats apparaissent encore moins discriminants, avec un  $R^2CV$  de 0,58 pour les spectres UV-visible et un  $R^2CV$  de 0,75 pour les spectres IRTF.

Si l'on compare ces résultats avec ceux obtenus pour le suivi de la maturité du raisin, nous constatons des résultats très proches pour les analyses de références, alors que pour les spectres UV-visible et IRTF, la discrimination baisse fortement. Rappelons que pour les spectres lors de la maturité, la prédiction du millésime était pratiquement parfaite, avec un  $R^2CV$  de 0,99 pour l'UV-visible et de 1,00 pour l'IRTF. Deux hypothèses peuvent expliquer ces résultats : premièrement, il est possible que le millésime n'impacte pas aussi fortement la matrice pendant la vinification que pendant la maturité. Cependant, les résultats très proches des analyses de références ne semblent pas corrélés cette hypothèse. La deuxième possibilité est que la variation de la composition de la matrice (sucre, éthanol...) ne permet pas de mettre en évidence les variations des tanins et anthocyanes dues aux millésimes. Ainsi, l'impact des composés phénoliques serait masqué par les évolutions des autres composés.

Dans l'ensemble, bien que les discriminations manquent de robustesse, nous ne pouvons pas exclure l'effet millésime sur les spectres UV-visible et IRTF, ce qui est à éviter afin d'obtenir des modèles de prédiction robustes. Il sera donc nécessaire de travailler sur des modèles de prédiction validés par un millésime extérieur. Concernant un possible effet cépage, les résultats obtenus par PLSDA nous permettent de conclure que les variétés utilisées influencent très peu les concentrations en tanins et anthocyanes mesurées. De plus, nous ne constatons pas d'empreinte spectrale propre à chaque cépage.

## **2. Evolution des concentrations en anthocyanes et en tanins pendant la fermentation alcoolique**

Avant de discuter des résultats de prédiction des différents paramètres étudiés, il est important d'évaluer l'évolution des anthocyanes et des tanins pendant la fermentation alcoolique. Nous avons ainsi pu suivre les concentrations en anthocyanes grâce à deux analyses différentes : la méthode de dosage au bisulfite de sodium et par séparation en CLHP-UV-visible. Pour compléter ces informations, nous nous sommes intéressés à la polymérisation de ces composés grâce au dosage des pigments monomériques et polymérisés décrit par Adams et Harbertson. Pour obtenir les concentrations en tanins, nous avons utilisé deux méthodes de dosage : la précipitation à l'albumine de sérum bovin (BSA) et la précipitation à la méthylcellulose (MCP). De plus, pour évaluer la

composition des unités flavan-3-ols qui composent les proanthocyanidines, nous avons analysé par CHLP/UV-visible le produit de la réaction des tanins avec le phloroglucinol.

### a) Evolution des concentrations en anthocyanes pendant la fermentation alcoolique

La Figure 63 présente le suivi des concentrations en anthocyanes libres (par décoloration au bisulfite) et la somme des anthocyanes moléculaires obtenue par CLHP-UV-visible.

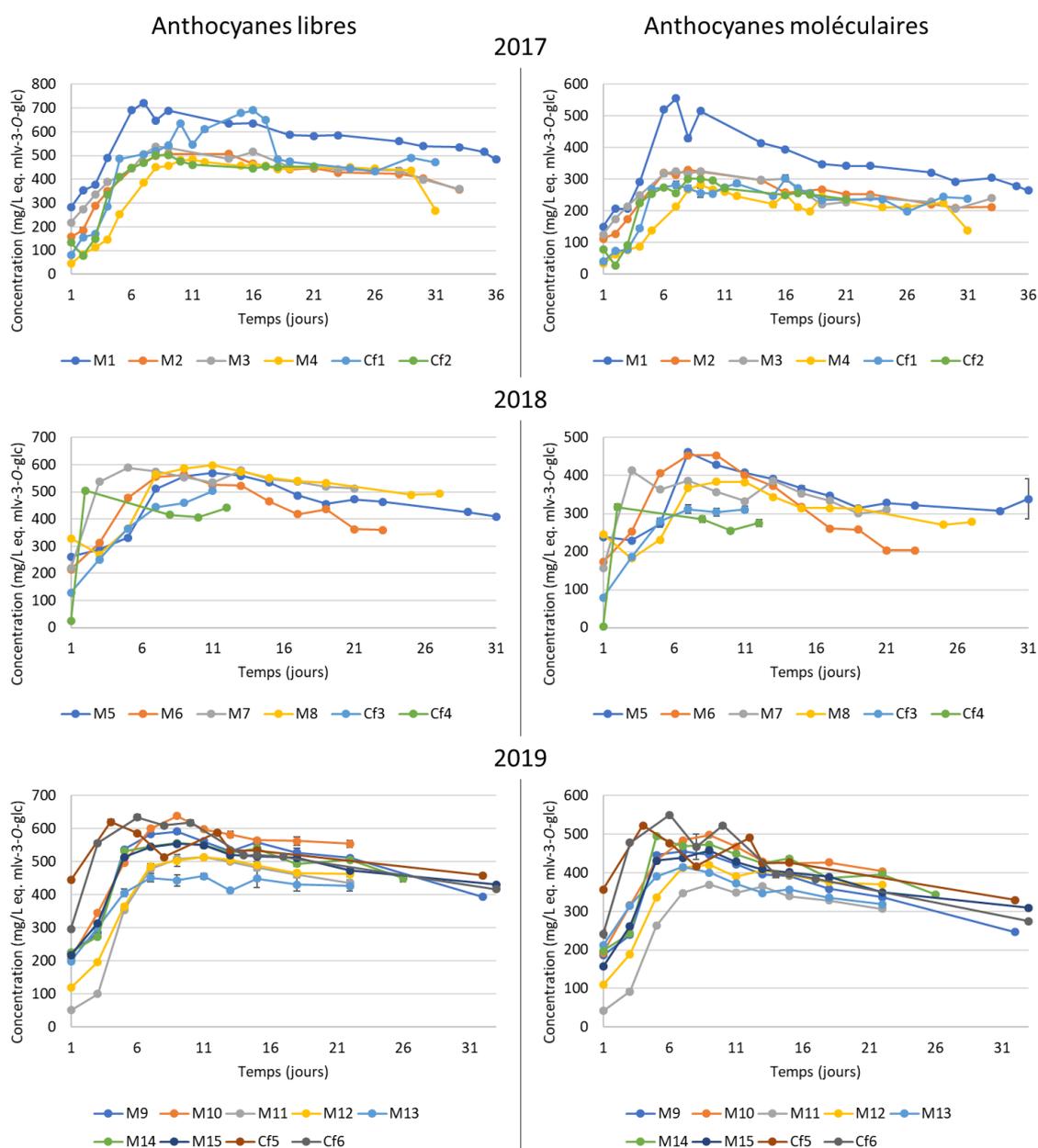


Figure 63 : Evolution de la concentration en anthocyanes libres et anthocyanes moléculaires sur les millésimes 2017, 2018 et 2019 (M : Merlot, Cf : Cabernet franc) (Mlv-3-O-glc : malvidine-3-O-glucoside)

Nous constatons via les deux méthodes d'analyse une évolution caractéristique de la concentration en anthocyanes. Ces résultats indiquent une augmentation rapide durant la première semaine de la fermentation alcoolique, suivie d'une diminution régulière jusqu'à l'écoulage. L'extraction des anthocyanes commence donc dès l'encuvage, avec des concentrations qui varient entre un minimum de 23,62 mg/L et un maximum de 414,15 mg/L pour les anthocyanes libres, et entre un minimum de 3,49 mg/L et un maximum de 356,83 mg/L pour les anthocyanes moléculaires. Les concentrations atteignent leur maximum après une semaine, où nous observons un maximum de 721,20 mg/L pour les anthocyanes libres et de 556,18 mg/L pour les anthocyanes moléculaires. Après ce pic, les concentrations commencent à décroître lentement et de manière constante jusqu'à l'écoulage, avec un minimum de 268,30 mg/L et un maximum de 554,34 mg/L pour les anthocyanes libres, et un minimum de 139,64 mg/L et un maximum de 404,20 mg/L pour les anthocyanes moléculaires. Cette évolution spécifique est expliquée par la bonne extractibilité des anthocyanes en milieu aqueux, facilitée par l'augmentation du volume d'éthanol (350). Premièrement, les anthocyanes sont rapidement extraites de la pellicule du raisin jusqu'à atteindre un maximum. Ensuite, une fois solubilisés, ces composés peuvent aussi se dégrader, interagir et polymériser avec les flavan-3-ols pour former des pigments oligomériques et polymériques (194,351). Les écarts entre les minimums et maximums sont importants, et montrent bien l'hétérogénéité des concentrations sur les différentes cuves et millésimes suivis. De plus, l'évolution spécifique des anthocyanes met en évidence l'importance de leur suivi, qui permet de définir la fin de leur extraction et de suivre leur transformation.

En comparant les deux méthodes d'analyse, nous constatons qu'elles permettent d'obtenir des résultats très similaires, mais certaines divergences sont observables. Nous remarquons que les concentrations mesurées ne sont pas identiques, avec des valeurs plus élevées pour les anthocyanes libres. Cette différence peut être attribuée à la méthode d'analyse et l'utilisation d'un étalonnage différent. Concernant l'évolution des différentes cuves, très peu de différences sont notables. Nous observons toutefois des divergences, comme par exemple pour le millésime 2018, où la cuve M8 apparaît être la plus concentrée en anthocyanes libres, alors qu'il s'agit de la cuve M5 pour les anthocyanes moléculaires. Rappelons que les anthocyanes libres sont mesurés grâce à une analyse colorimétrique par décoloration au bisulfite de sodium. Ainsi, un facteur comme la copigmentation peut fortement influencer la lecture. La valeur d'anthocyanes libres est donc un bon indicateur de la couleur induite par les anthocyanes non polymérisées, alors que la valeur d'anthocyanes moléculaires permet une vision plus spécifique de la composition moléculaire.

## ***b) Evolution de la composition en anthocyanes moléculaires pendant la fermentation alcoolique***

Nous avons pu constater que les anthocyanes avaient une évolution spécifique lors de la fermentation. Il est désormais nécessaire d'étudier la proportion de chaque anthocyane moléculaire lors du processus de vinification, afin d'évaluer si elles ont une évolution distincte les unes des autres ou si elles suivent le même schéma d'extraction et d'évolution. Pour cela, nous avons étudié la distribution des différentes anthocyanes moléculaires en fonction des cuves étudiées lors de la vinification (Figure 64). L'écart-type présenté correspond à la variabilité de chaque anthocyane moléculaire pour tous les prélèvements d'une même cuve.

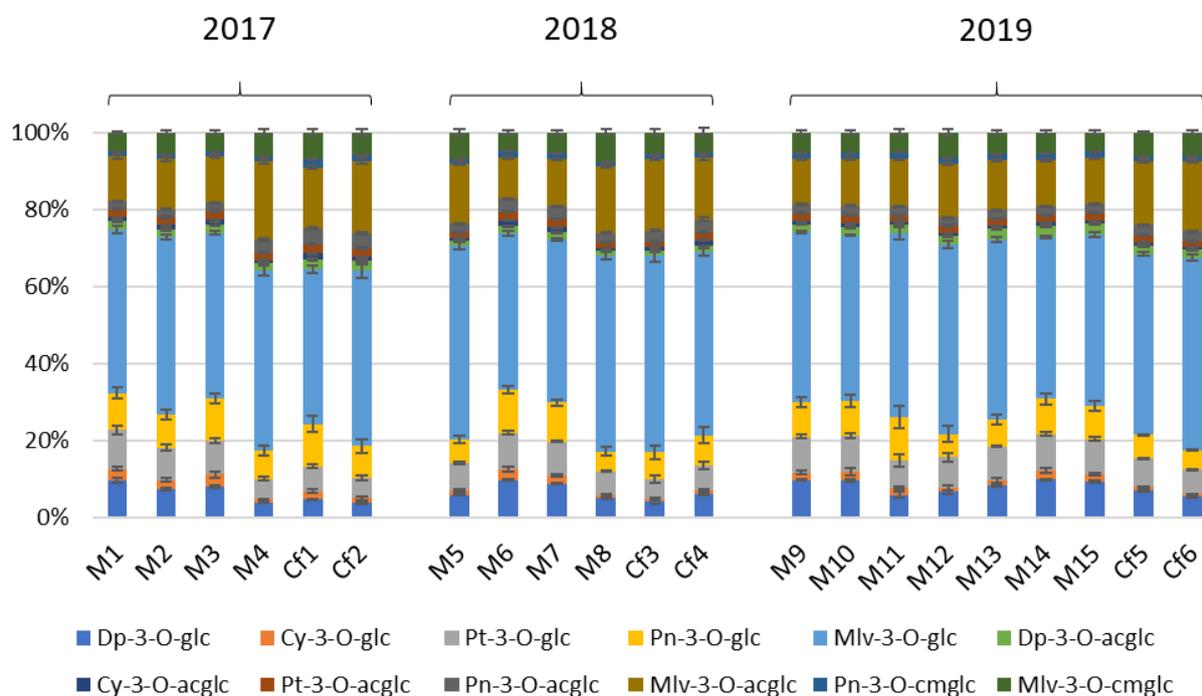


Figure 64 : Distribution des différentes anthocyanes moléculaires en fonction des cuves de vinification étudiées pour les millésimes 2017, 2018 et 2019 (M : Merlot, Cf : Cabernet franc) (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside)

Nous constatons que la répartition des anthocyanes est sensiblement identique à celle obtenue lors du suivi de la maturité du raisin. Pour 2018, il est possible de comparer directement les résultats de la fermentation avec ceux de la maturité du raisin (Figure 52), car pour ce millésime les cuves de vinification suivies correspondent aux parcelles étudiées. Des différences sont alors notables : nous remarquons par exemple que le taux de cyanidine-3-O-glucoside apparaît inférieur à celui obtenu sur le raisin. Pendant le suivi de maturité en 2018, la proportion moyenne de cyanidine-3-O-glucoside

était de 3,3 % avec un maximum de 6 %, alors que pendant la fermentation, la proportion moyenne de cyanidine-3-*O*-glucoside est de 1,4 % avec un maximum de 2,8 %. Pour les autres composés, les proportions apparaissent très proches, avec par exemple un taux moyen de malvidine-3-*O*-glucoside (l'anthocyane majoritaire), de 47,1 % pour la maturité et la fermentation.

Concernant l'évolution des proportions pendant la fermentation, nous observons comme pour la maturité un écart-type très faible. Chacune des différentes anthocyanes semblent donc suivre le même schéma d'extraction. Afin de mettre clairement en évidence cette constatation, intéressons-nous à la cuve M11, qui montre le plus grand écart-type pour la plupart des composés, et notamment pour la malvidine-3-*O*-glucoside et la péonidine-3-*O*-glucoside. Sur la Figure 65 on peut suivre l'évolution de la concentration des différentes anthocyanes moléculaires durant la fermentation alcoolique pour cette cuve, ainsi que leur distribution exprimée en pourcentage.

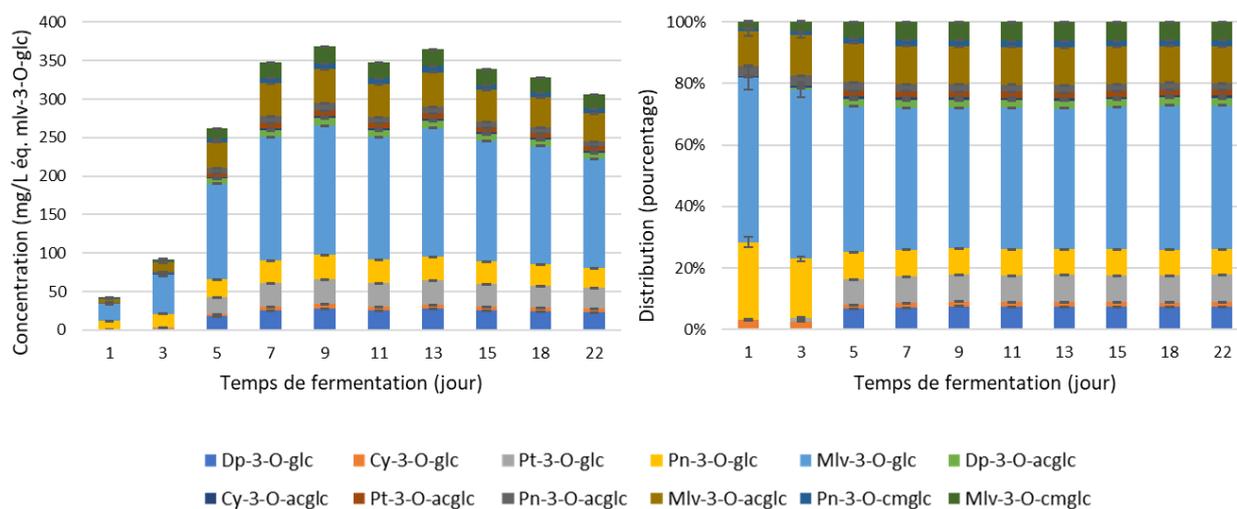


Figure 65 : Evolution de la concentration et la distribution des anthocyanes moléculaires de la cuve M11 durant la fermentation alcoolique (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Malv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside)

Nous observons pour cette cuve une extraction rapide jusqu'à un maximum suivie d'une diminution lente de la concentration globale jusqu'à l'écouage. Nous remarquons que pour les deux premiers prélèvements, la distribution des anthocyanes apparaît différente des autres temps de la fermentation, avec un pourcentage en malvidine-3-*O*-glucoside et péonidine-3-*O*-glucoside plus important. Cette distinction explique l'écart-type important observé sur la Figure 64, et ces composés pourraient être extraits préférentiellement en début de fermentation. Cependant, à partir du cinquième jour de fermentation et jusqu'à l'écouage, la distribution des anthocyanes est identique, et ce malgré l'évolution de la concentration globale en anthocyanes. La concentration de chaque anthocyane se

retrouve donc fortement corrélée avec la concentration en anthocyanes totales, et cette donnée devra être prise en compte lors de la conception des modèles de prédiction pour éviter une mauvaise interprétation des résultats. Néanmoins, nous avons pu constater que la distribution d'une cuve à l'autre n'était pas identique (Figure 64), il reste donc intéressant de travailler sur la prédiction des anthocyanes moléculaires pendant la fermentation pour pouvoir tester l'utilité et les limites de l'UV-visible et l'IRTF dans le dosage de ces composés.

### *c) Evolution de la polymérisation des anthocyanes pendant la fermentation alcoolique*

Nous avons pu évaluer l'évolution de la concentration en anthocyanes libres ainsi que la distribution des anthocyanes moléculaires durant la fermentation. Si l'on s'intéresse à l'évolution de leur polymérisation, grâce à la méthode de dosage Adams-Harbertson, nous obtenons l'abondance relative des pigments monomériques, des petits pigments polymérisés (les procyanidines possédant moins de 4 unités flavan-3-ols et ne pouvant se lier avec l'albumine de sérum bovin) et des grands pigments polymérisés (275). Ces proportions exprimées en pourcentage ont pu être suivies pour les différentes cuves et les 3 millésimes étudiés, en fonction du temps de fermentation (Figure 66).

La première observation notable est la répartition très divergente pour certaines cuves suivies lors des premiers jours de vinification. Si l'on s'intéresse à la cuve M2, Cf4 et M12, nous constatons un ratio important de pigments polymérisés. Pour la cuve M2, les grands pigments polymérisés composent plus de 40 % des anthocyanes le premier jour, pour drastiquement diminuer en dessous de 20 % le deuxième jour et ce jusqu'à l'écoulage. La même observation peut être faite pour les cuves Cf4 et M12. Cependant, il faut rappeler que le premier jour de fermentation, les concentrations en anthocyanes restent faibles (Figure 63). Les valeurs obtenues lors du début de la vinification peuvent donc être exclues car la méthode de dosage utilisée peut manquer de précision avec des concentrations faibles. Cette hypothèse est corroborée par les écarts type plus importants observés sur le début de la fermentation. Néanmoins, ces divergences ne sont observables que pour les premiers jours, et n'entravent pas la pertinence de la méthode de dosage pour suivre l'évolution des anthocyanes lorsque celles-ci sont extraites de la baie de raisin en plus grande quantité.

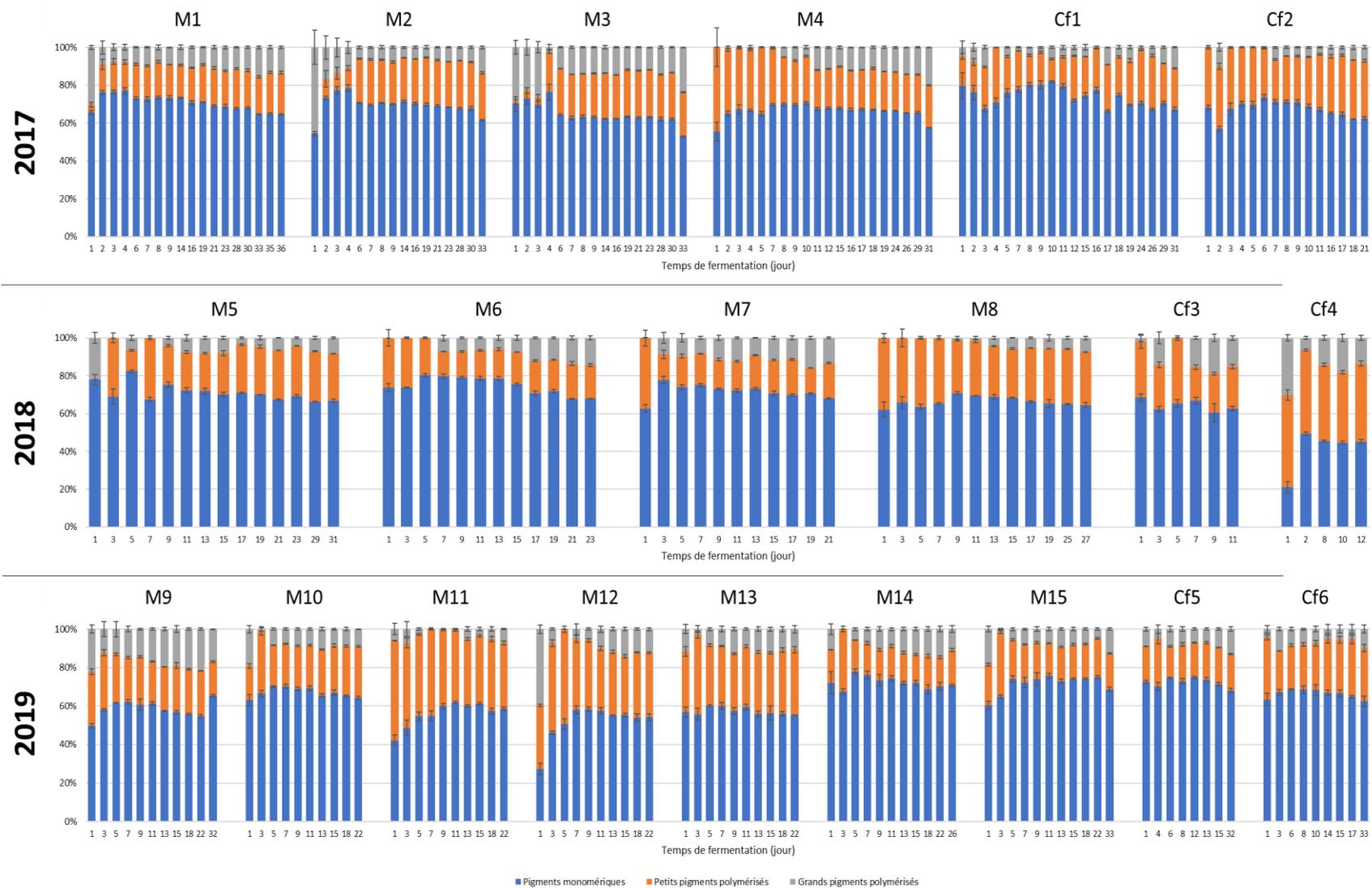


Figure 66 : Evolution de la polymérisation des anthocyanes en fonction du temps de fermentation pour les millésimes 2017, 2018 et 2019 sur les vins en fermentation (M : Merlot, Cf : Cabernet franc)

Après les premiers jours de fermentation, les proportions des différents pigments suivent une évolution caractéristique qui est observable sur toutes les cuves étudiées et sur tous les millésimes. Au moment du pic de concentration des anthocyanes moléculaires, c'est-à-dire après une semaine de vinification (Figure 63), les pigments monomériques sont majoritaires. À partir de ce maximum, leurs proportions vont décroître de manière constante au profit d'une augmentation des pigments polymérisés. Si on prend l'exemple de la cuve M1, les pigments monomériques passent de 76 % au deuxième jour de fermentation à 64 % le jour de l'écouage. Quant aux pigments polymérisés, les petits pigments passent de 14 à 22 %, alors que les grands pigments passent de 9 à 13 %. Cette évolution de la polymérisation est observable sur toutes les cuves suivies, avec des divergences notables dans les proportions finales en pigments monomériques et polymérisés le jour de l'écouage. Ainsi, si nous comparons les cuves M7 et M13, qui partagent un temps de fermentation et macération très proche (21 et 22 jours), nous remarquons des proportions en pigments le jour de l'écouage très différentes. La cuve M7 présente une proportion en pigments monomériques en fin de macération de 68 %, et la cuve M13 de 55 %. Les petits et grands pigments polymérisés représentent respectivement 19 et 13 % pour la cuve M7, et 34 et 11 % pour la cuve M13. Ces deux cuves sont intéressantes à comparer car si l'on regarde la concentration en anthocyanes moléculaires (Figure 63) nous remarquons que leurs concentrations le jour de l'écouage sont proches, aux alentours de 300 mg/L. La polymérisation des anthocyanes est donc indépendante d'une cuve à l'autre, et il est important de pouvoir la suivre afin d'évaluer son évolution tout au long de la fermentation.

La concentration et la polymérisation des anthocyanes jouent un rôle prépondérant sur la couleur des vins. Nous avons pu constater que lors de la fermentation et de la macération post-fermentaire, ces anthocyanes évoluent, et montrent des concentrations très différentes d'une cuve à l'autre. De plus, certains processus de vinification peuvent impacter cette évolution, tel que l'ajout de tanins ou la micro-oxygénation (158,352,353). Il est donc nécessaire pour le vinificateur de pouvoir suivre correctement ces molécules. Le jeu de données utilisé montre bien les différentes évolutions des anthocyanes durant la vinification et peuvent permettre de développer des modèles de régression PLS pour IRTF et UV-visible efficaces pour leur prédiction.

### d) Evolution des concentrations en tanins pendant la fermentation alcoolique

La Figure 67 présente le suivi des concentrations en tanins dosés par précipitation à l'albumine de sérum bovin (BSA) et par précipitation à la méthylcellulose (MCP).

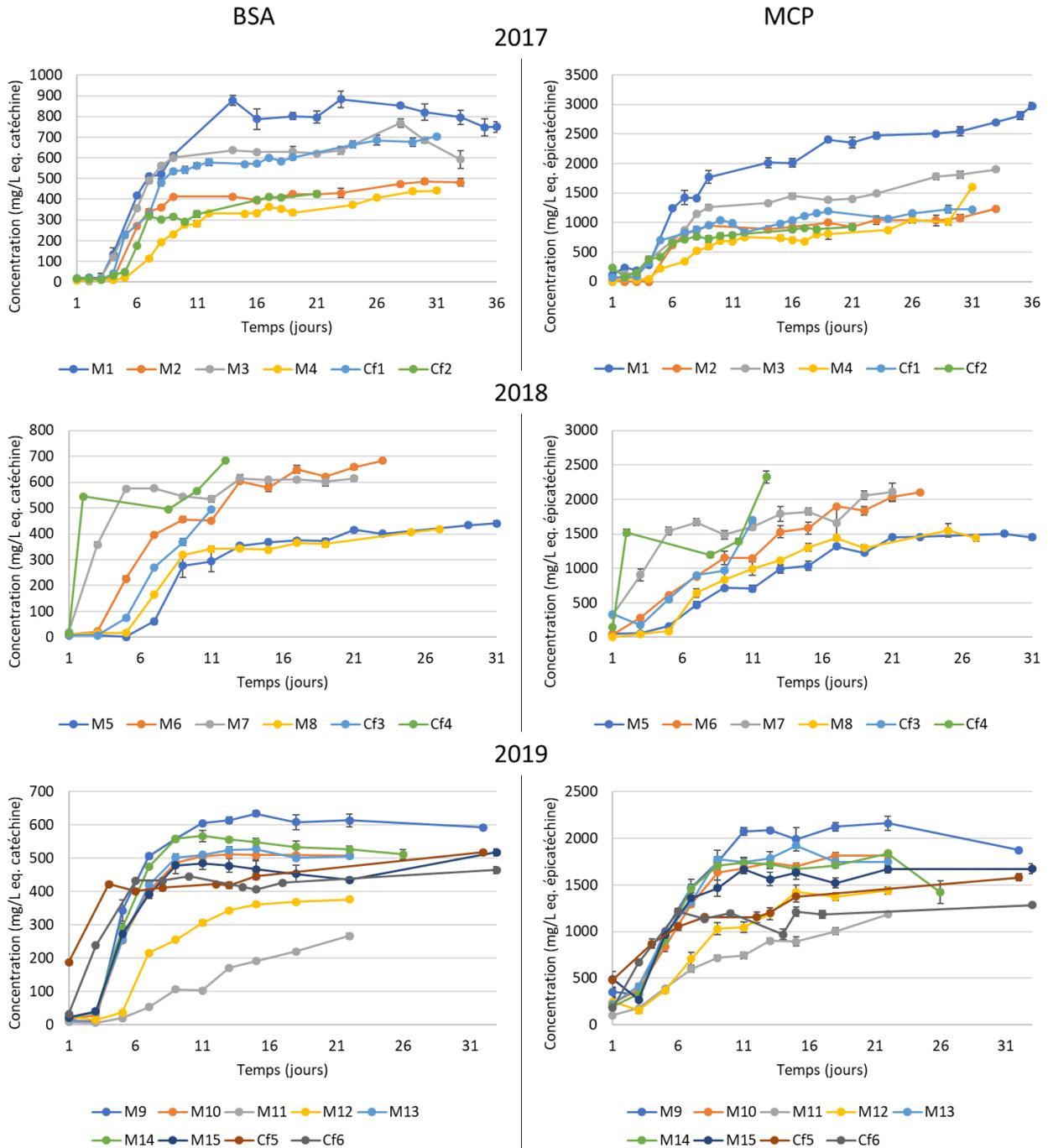


Figure 67 : Evolution des concentrations en tanins dosés par précipitation à l'albumine de sérum bovin (BSA) et par précipitation à la méthylcellulose (MCP) en fonction du temps de fermentation pour les cuvées des millésimes 2017, 2018 et 2019 (M : Merlot, Cf : Cabernet franc)

Pour les trois millésimes suivis, nous observons une évolution de la concentration en tanins très proche. Les premiers jours de vinification, qui correspondent au temps de macération pré-fermentaire à froid, montrent des teneurs en tanins faibles. Ces composés ont une hydrosolubilité limitée et l'absence d'éthanol dans le milieu ne permet pas une extraction idéale. Lorsque la fermentation commence et que le taux en éthanol augmente, nous observons une accélération de l'extraction pendant 2 à 7 jours. Ensuite, pendant la macération post-fermentaire, l'extraction continue de manière plus lente et constante. Même si ce schéma d'évolution est applicable à la majorité des cuves suivies, nous observons cependant des différences d'une cuve à l'autre, notamment pendant la macération post-fermentaire. Prenons l'exemple du millésime 2018 et des cuves M6 et M7 avec des concentrations en tanins proches le jour de l'écoulage, avec respectivement 682,99 et 614,40 mg/L équivalent catéchine pour BSA et 2105,69 et 2109,24 mg/L équivalent épicatechine pour MCP. L'extraction des tanins de la cuve M7 est très importante dès les premiers jours de la vinification, et la concentration atteint 574,66 mg/L équivalent catéchine pour BSA et 1541,53 mg/L équivalent épicatechine pour MCP après seulement 5 jours. A contrario, pour la cuve M6, l'extraction est plus lente, et après le même nombre de jours, la concentration en tanins est seulement de 226,82 mg/L équivalent catéchine pour BSA et 612,05 mg/L équivalent épicatechine pour MCP. L'extraction des tanins est donc indépendante d'une cuve à l'autre, et cet exemple montre tout l'intérêt pour le vinificateur de pouvoir suivre rapidement et facilement ces teneurs, afin d'adapter au mieux les différents processus tels que le remontage, pigeage ou date d'écoulage pour obtenir la concentration en tanins recherchée.

Si les deux méthodes de dosage cherchent à mesurer les tanins, nous constatons cependant certaines différences entre les deux évolutions. La méthode de dosage à l'albumine de sérum bovin (BSA) différencie plus fortement les cuves suivies que la méthode à la méthylcellulose (MCP). Pour le millésime 2017, nous observons pour le paramètre BSA que la cuve M1 se démarque des autres par une concentration en tanins précipitables supérieure (800 mg/L équivalent catéchine), suivie par les cuves M3 et Cf1 qui partagent des teneurs sensiblement identiques (600 mg/L équivalent catéchine). Les trois autres cuves suivies conduisent à des concentrations similaires proches de 400 mg/L équivalent catéchine. Pour la méthode MCP, nous remarquons aussi un maximum pour la cuve M1 (jusqu'à 3500 mg/L équivalent épicatechine), qui se démarque fortement des autres cuves (maximum de concentration compris entre 1500 et 2000 mg/L équivalent épicatechine), mais les différences entre les cinq cuves restantes ne sont pas aussi marquées que pour le paramètre BSA. Les mêmes observations peuvent être faites pour le millésime 2018, où nous constatons une forte différence pour

BSA entre les cuves M6, M7, Cf3, Cf4 (maximum de concentration compris entre 600 et 700 mg/L équivalent catéchine) et les cuves M5, M8 (avec un maximum de concentration compris entre 300 et 350 mg/L équivalent catéchine), alors que cette distinction n'est pas visible pour MCP. Pour le millésime 2019, ce sont les cuves M11 et M12 qui se démarquent avec une concentration en tanins inférieure pour BSA, alors que pour MCP, seule la cuve M11 conduit à des teneurs en tanins légèrement inférieures au reste du suivi. Ces variations mettent en évidence les différences entre les deux dosages, avec la méthode de dosage à l'albumine de sérum bovin qui précipitera préférentiellement les tanins composés de 4 unités flavan-3-ols ou moins, alors que le dosage à la méthylcellulose n'a pas cette spécificité. Tout comme pour le suivi de la maturité du raisin, les deux méthodes apportent une information différente, et leur comparaison permet non seulement de suivre l'évolution de l'extraction des tanins, mais aussi de juger de leur réactivité.

#### ***e) Evolution de la composition en unités flavan-3-ols des tanins pendant la fermentation alcoolique***

Après avoir étudié l'évolution de la concentration en tanins, intéressons-nous à leur composition en unités flavan-3-ols. Nous avons obtenu la concentration en unités catéchine, épicatechine, épigallocatechine et epicatechine-gallate grâce à la méthode de dépolymérisation par réaction au phloroglucinol. La Figure 68 présente le pourcentage de chaque monomère en fonction du temps de fermentation. Globalement, nous constatons que les proportions des flavan-3-ols n'évoluent que très peu durant la majeure partie de la vinification. Pour toutes les cuves étudiées, l'épicatechine est majoritaire et l'épicatechine-gallate minoritaire. Cependant, certaines différences pour les composés sont observables. En début de fermentation, nous remarquons que sur plusieurs des cuves suivies, l'épigallocatechine est majoritaire, et son abondance diminue progressivement au profit de l'épicatechine. Nous observons distinctement cette évolution caractéristique sur les cuves M1, M2, M3, M5, M11 et M12. L'épigallocatechine est un marqueur de la pellicule de raisin, et ne se retrouve pas dans les pépins. Nous pouvons donc supposer que cette forte proportion en début de fermentation est due à l'extraction des tanins de la pellicule qui se solubilisent plus rapidement et facilement que ceux des pépins (109,110).

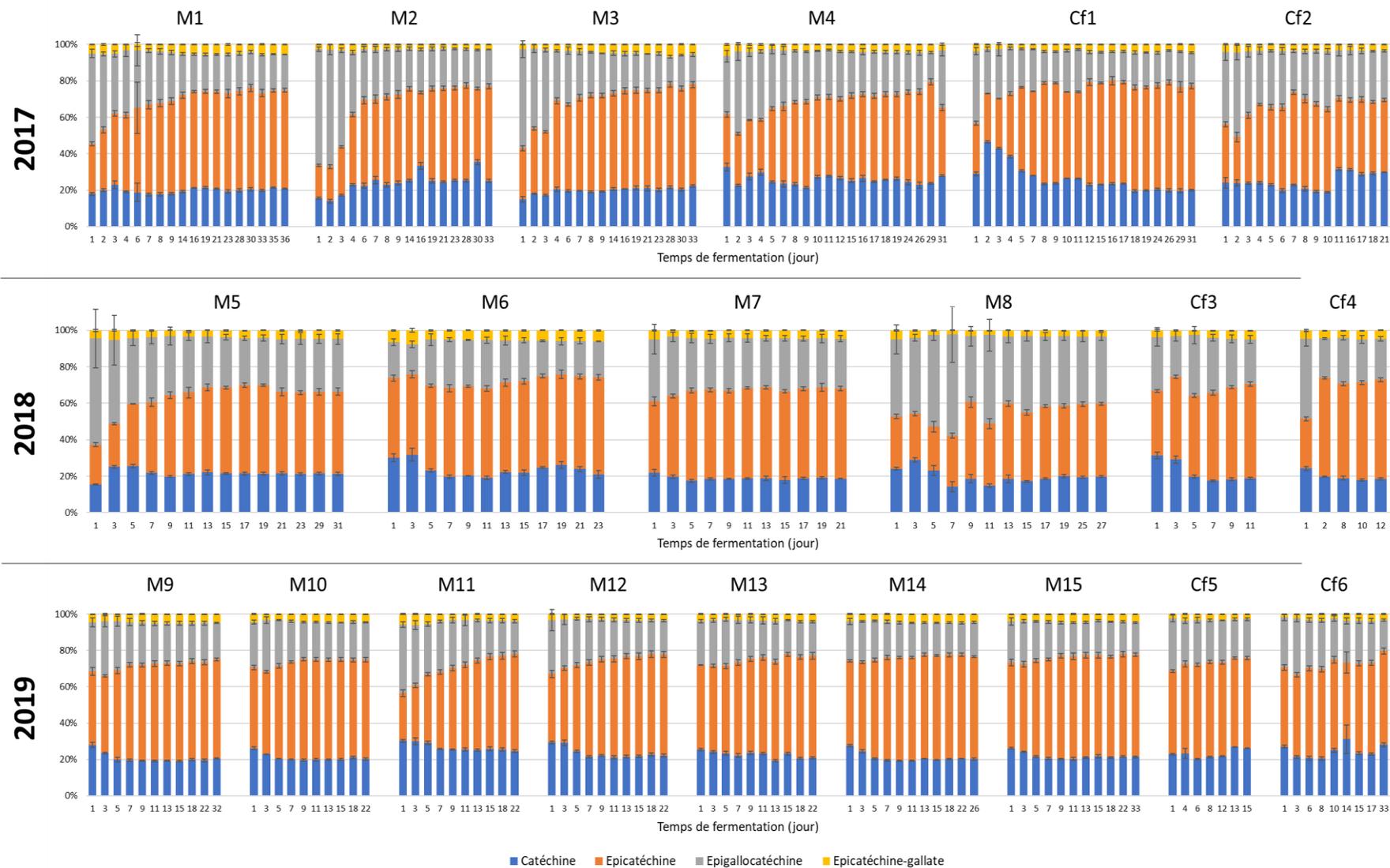


Figure 68 : Evolution de la composition en unités flavan-3-ols en fonction du temps de fermentation pour les millésimes 2017, 2018 et 2019 sur les vins en fermentation (M : Merlot, Cf : Cabernet franc)

Nous distinguons également des disparités dans les proportions de chaque monomère tannique en fonction des cuves étudiées. La catéchine apparaît comme le flavan-3-ol avec l'abondance la plus constante, proche de 20 %. Nous observons toutefois des différences notables : ainsi, le jour de l'écouage, les cuves Cf2 et Cf6 montrent respectivement des proportions en catéchine de 29,9 % et 28,11 %, alors que pour les cuves M7 et Cf4 ce pourcentage est de 18,5 %. De plus, nous remarquons que pour plusieurs cuves (Cf1, Cf3 M9, M10, M11, M12, M14, M15), ces proportions en début de fermentation sont supérieures au reste de la vinification, pouvant atteindre jusqu'à 46,5 % le deuxième jour de fermentation pour Cf1. Ces différences entre cuves apparaissent encore plus importantes lorsque nous étudions les autres monomères. Ainsi, en comparant la cuve M1 et M8, nous constatons de forts écarts dans la répartition des flavan-3-ols. Pour la cuve M1 le jour de l'écouage, le pourcentage de catéchine est de 20,8 %, de 54,0 % pour l'épicatéchine, de 19,6 % pour l'épigallocatechine, et de 5,6 % pour l'épicatéchine-gallate. Pour la cuve M8, le pourcentage de catéchine est de 19,8 %, de 39,9 % pour l'épicatéchine, de 37,0 % pour l'épigallocatechine, et de 3,3 % pour l'épicatéchine-gallate. Si les proportions en catéchine sont pratiquement identiques pour les deux cuves, pour les autres monomères, les différences sont importantes, notamment pour l'épigallocatechine où la cuve M8 présente une abondance proche du double de celle de la cuve M1, ou pour l'épicatéchine-gallate où les proportions de la cuve M1 sont bien supérieures à la cuve M8. L'épicatéchine-gallate est un marqueur des tanins du pépin de raisin, a contrario de l'épigallocatechine qui ne se trouve que dans la pellicule (109,111). En étudiant ces abondances, nous pouvons donc observer que l'extraction des tanins de pépins est plus importante pour la cuve M1 que pour M8. Ce type d'information peut s'avérer utile au vinificateur afin de s'assurer du bon déroulement de l'extraction des tanins durant la fermentation alcoolique et la macération post-fermentaire. Il est donc particulièrement intéressant d'essayer de modéliser les proportions des différents flavan-3-ols. De plus, les résultats observés entre les différentes cuves du jeu de données montrent une bonne variabilité, et semblent tout à fait adaptés à la conception de modèles de prédiction.

### ***f) Evolution du degré de polymérisation moyen des tanins pendant la fermentation alcoolique***

Pour parfaire les informations obtenues sur les tanins pendant la fermentation alcoolique, nous avons pu calculer le degré de polymérisation moyen grâce au ratio entre les concentrations en unités flavan-3-ols terminales et totales obtenues après la dépolymérisation par

réaction au phloroglucinol. La Figure 69 présente le degré de polymérisation moyen des différentes cuves suivies en fonction du temps de fermentation.

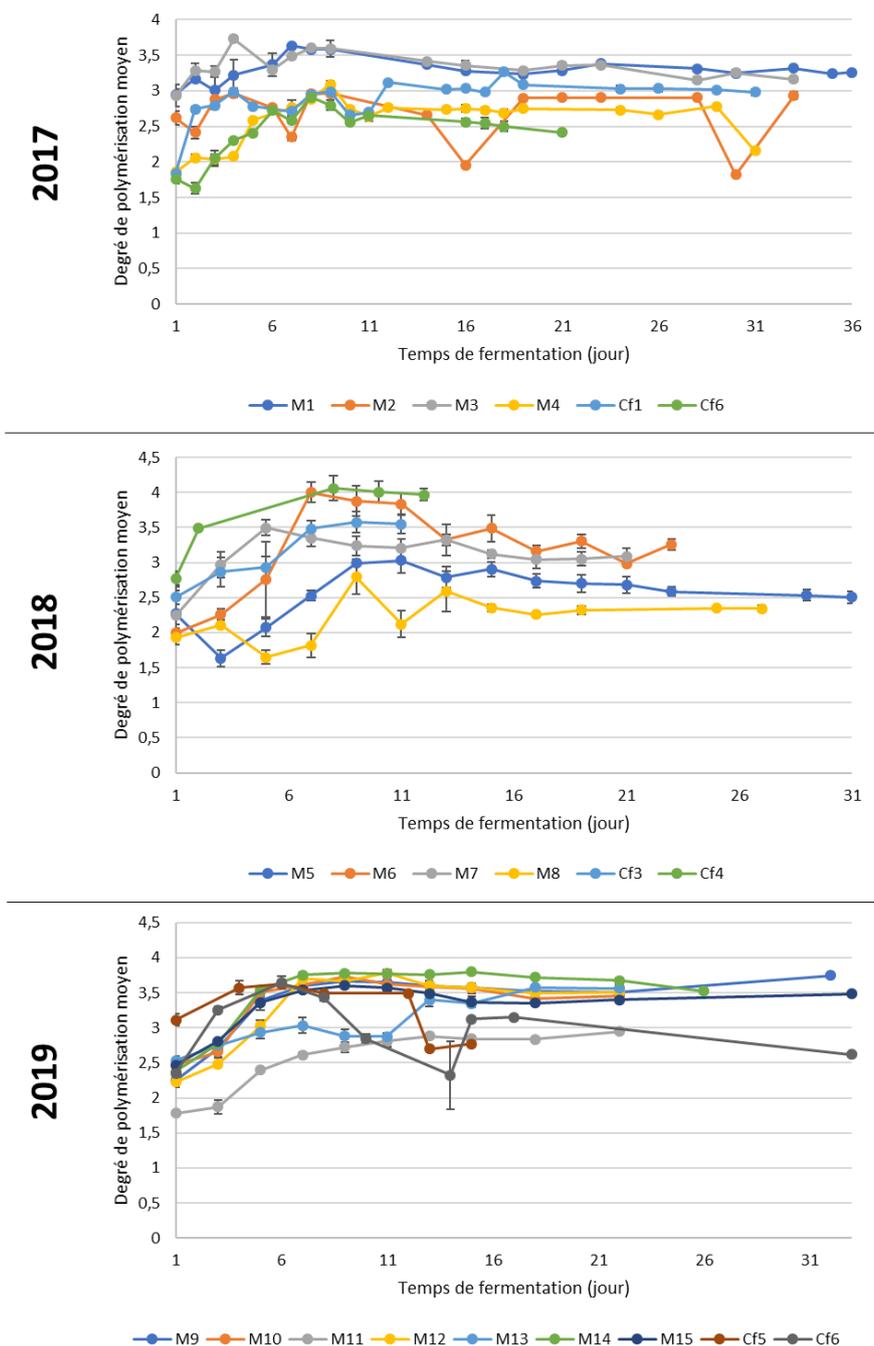


Figure 69 : Evolution du degré de polymérisation moyen en fonction du temps de fermentation pour les millésimes 2017, 2018 et 2019 pour les vins en fermentation (M : Merlot, Cf : Cabernet franc)

Pour la majorité des cuves suivies, le degré de polymérisation moyen (DPM) est à son minimum le premier jour de la vinification. Il augmente graduellement lors de la première semaine, pour atteindre un maximum qui n'évoluera que très peu pendant le reste de la fermentation. Nous constatons

cependant une diminution progressive pour la plupart des cuves, qui est distinctement observable sur les cuves M5 et M6 du millésime 2018. On peut également remarquer que la variation du DPm n'est pas élevée, et la grande majorité des données recueillies sont comprises entre 2,5 et 3,5. Si l'on calcule le coefficient de variation, qui est défini par le rapport entre l'écart-type et la moyenne, nous obtenons un résultat de 17 %. Celui-ci apparaît très inférieur à ceux obtenus pour les autres paramètres étudiés, le plus faible étant relevé pour la concentration en anthocyanes libres avec 30 % (Tableau 21). Nous observons également des écarts-types plus importants que pour la plupart des autres analyses de références, qui sont notamment visibles sur le millésime 2018. Cependant, plusieurs cuves sont identifiables par un DPm plus élevé ou plus faible que la majorité du suivi, c'est notamment le cas des cuves M8 et Cf4 du millésime 2018 avec un DPm respectif le jour de l'écoulage de 2,34 et 3,97, soit un écart de 1,63. Cette différenciation apporte une information complémentaire dans l'analyse des tanins.

Au regard du manque de variabilité présenté par le suivi du DPm, sa modélisation par PLS doit être sujet à discussion. Nous observons suffisamment de différences pour que l'information apportée soit utile à l'appréciation des tanins, mais nécessite cependant que le modèle soit extrêmement robuste pour que l'erreur de prédiction soit minime en comparaison de la variabilité du jeu de données. De plus, cette valeur ne peut être modélisée directement, et doit être calculée comme le ratio entre les concentrations en unités flavan-3-ols terminales et totales. Les modèles correspondant à ces deux concentrations se doivent donc d'être prédictifs pour obtenir un résultat exploitable.

### **3. Construction de modèles de prédiction PLS pour le suivi des anthocyanes et des tanins pendant la vinification**

Nous avons étudié les concentrations et la variabilité des anthocyanes et des tanins des échantillons analysés pendant la vinification. Nous avons pu détailler leur extraction durant les différentes étapes de la fermentation, à partir du premier jour d'encuvage et jusqu'à l'écoulage. Nous avons mis en évidence l'évolution distincte des différents composés étudiés, ainsi que la variabilité recherchée dans le jeu de données pour permettre la conception de modèles de prédiction efficaces et représentatifs de la période de vinification. Nous allons désormais aborder les résultats des modèles de prédiction, en présentant dans un premier temps les modèles obtenus sur tous les échantillons en utilisant la validation croisée comme méthode de validation, puis ensuite les modèles obtenus par validation avec un jeu d'échantillons extérieurs à la calibration.

***a) Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification***

Grâce aux données des analyses de références recueillies pendant la fermentation, nous avons pu construire des modèles de prédiction en se basant sur les spectres IRTF, UV-visible et sur la somme des deux zones spectrales. Pour valider la modélisation, nous avons utilisé la validation croisée *leave-p-out cross-validation*, avec 10 sous-groupes. Pour définir le partitionnage des groupes, nous avons utilisé la méthode dite *venetian blinds*.

***(1) Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification par IRTF***

Le Tableau 23 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres IRTF.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL
<b>Composition en flavan-3-ols</b>						
Flavan-3-ol (mmol/L)	0,9	2,13	0,89	2,29	2,96	10
Flavan-3-ol terminal (mmol/L)	0,80	0,99	0,78	1,05	2,04	11
Catéchine (mmol/L)	0,85	0,52	0,83	0,55	2,40	9
Epicatéchine (mmol/L)	0,92	1,06	0,90	1,16	3,22	10
Epigallocatechine (mmol/L)	0,76	0,89	0,73	0,93	1,87	10
Epicatéchine-gallate (mmol/L)	0,86	0,14	0,83	0,15	2,40	9
<b>Anthocyanes moléculaires</b>						
Dp-3-O-glc (mg/L)	0,88	4,57	0,86	4,92	2,63	12
Cy-3-O-glc (mg/L)	0,73	1,52	0,66	1,70	1,69	9
Pt-3-O-glc (mg/L)	0,86	4,59	0,84	4,90	2,51	12
Pn-3-O-glc (mg/L)	0,69	4,99	0,61	5,59	1,53	13
Mlv-3-O-glc (mg/L)	0,90	15,52	0,88	17,16	2,86	13
Dp-3-O-acglc (mg/L)	0,62	2,11	0,59	2,19	1,55	4
Cy-3-O-acglc (mg/L)	0,61	0,78	0,58	0,81	1,44	6
Pt-3-O-acglc (mg/L)	0,79	1,18	0,75	1,27	1,98	8
Pn-3-O-acglc (mg/L)	0,68	1,69	0,64	1,84	1,61	8
Mlv-3-O-acglc (mg/L)	0,86	6,10	0,82	6,83	2,35	13
Pn-3-O-cmglc (mg/L)	0,86	0,86	0,81	0,97	2,31	13
Mlv-3-O-cmglc (mg/L)	0,89	2,70	0,87	2,98	2,73	13
Total glucoside (mg/L)	0,89	25,16	0,87	27,97	2,77	13
Total acétyl (mg/L)	0,86	8,83	0,82	9,97	2,37	13
Total coumaroyl (mg/L)	0,89	3,35	0,87	3,71	2,71	13
Total anthocyanes (mg/L)	0,89	35,62	0,87	39,62	2,72	13
<b>Evolution des anthocyanes</b>						
MP (Absorbance)	0,88	0,23	0,86	0,24	2,71	10
SPP (Absorbance)	0,76	0,11	0,74	0,11	1,95	7
LPP (Absorbance)	0,71	0,09	0,68	0,09	1,79	8
PP (Absorbance)	0,83	0,14	0,81	0,14	2,36	8
Anthocyanes libres (mg/L)	0,89	44,10	0,87	48,25	2,80	12
<b>Tanins</b>						
BSA (mg/L)	0,94	54,36	0,93	59,06	3,90	9
MCP (mg/L)	0,93	172,87	0,92	178,15	3,65	9

Tableau 23 : Modèles de prédiction PLS par IRTF avec validation croisée pour tous les paramètres d'anthocyanes et de tanins suivis pendant la fermentation (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Globalement, nous constatons que beaucoup de paramètres étudiés montrent des résultats de prédiction robustes qui pourraient être utilisés dans l'industrie. Si l'on s'intéresse à la composition en flavan-3-ols, nous remarquons que hormis pour l'épigallocatechine, tous les modèles indiquent un RPD CV (écart prédictif résiduel de validation croisée supérieur à 2). Rappelons que cette valeur, calculée comme le ratio entre la déviation standard du jeu de données et l'erreur quadratique moyenne de la validation croisée (RMSECV) permet de juger de la performance d'un modèle, et qu'au-dessus de 2 un modèle est considéré comme prédictif et qu'au-dessus de 3 il est considéré comme excellent (6,10,354). En regardant les deux paramètres qui permettent de calculer le degré de polymérisation moyen (DPM), nous remarquons que pour la concentration en unités flavan-3-ols totales (Flavan-3-ol), la prédiction apparaît robuste, avec un RPD CV de 2,96 et un  $R^2$ CV de 0,89. Ce paramètre peut être utilisé seul car il permet de doser les unités flavan-3-ols, et donc par extension les tanins, sans tenir compte de leur degré de polymérisation. Le deuxième paramètre utilisé pour calculer le DPM est la concentration en unités flavanols terminales (Flavan-3-ol terminal). Si les résultats démontrés par les modèles sont encourageants (RPD CV de 2,04 et un  $R^2$ CV de 0,78), ils ne sont pas suffisamment robustes pour être considérés. En effet, compte tenu de la faible variabilité du DPM observé sur notre jeu de données, son calcul grâce à nos modèles de prédiction ne sera pas assez précis pour distinguer les faibles variations observées. Dans la littérature, une seule étude à ce jour a cherché à déterminer le DPM des vins en utilisant l'analyse PLS couplée à l'infrarouge (288). Les résultats avancés sont très intéressants ( $R^2$  de 0,95) mais ont été obtenus sur des échantillons préalablement purifiés, non adaptés à l'analyse rapide. De plus, le fait que la prédiction soit construite directement sur la valeur de DPM est discutable. En effet, nous utilisons la spectroscopie pour la prédiction des anthocyanes et des tanins car la loi de Beer-Lambert s'applique et permet ainsi d'obtenir une réponse spectrale proportionnelle à la concentration. Il apparaît ainsi difficile d'obtenir un résultat fiable en essayant de modéliser un ratio tel que le DPM, sans risquer d'avoir une prédiction issue d'une interprétation erronée du jeu de données. Concernant les concentrations en différentes unités flavan-3-ols, nous remarquons de bons résultats, notamment pour l'épicatéchine avec un RPD CV de 3,22. La catéchine et l'épicatéchine-gallate montrent des prédictions similaires, avec un RPD CV de 2,40. Ces prédictions sont très intéressantes, notamment pour l'épicatéchine-gallate, car pouvoir suivre l'évolution de sa concentration permet d'estimer la proportion de tanins de pépins extraits.

Pour la prédiction des anthocyanes moléculaires, les modèles élaborés sont très proches de ceux obtenus pour le suivi de la maturité du Tableau 15. Nous constatons de bons résultats pour la majorité des composés, et une bonne prédiction de la somme des anthocyanes moléculaires (total anthocyanes)

avec un RPD CV de 2,72, ainsi que pour la somme des anthocyanes glucosides (RPD CV de 2,77), acétylglucosides (RPD CV de 2,37) et coumaroylglucosides (RPD CV de 2,71). Nous avons pu observer sur l'analyse du raisin un biais de prédiction dû à la forte corrélation entre la plupart des anthocyanes moléculaires et la somme de ces anthocyanes. Ainsi, il est nécessaire de directement infirmer ou confirmer si ce biais est observable sur le suivi de la fermentation. Pour cela, nous avons étudié cette corrélation et nous avons essayé de prédire chaque anthocyane grâce au modèle obtenu pour la somme des anthocyanes moléculaires. Les résultats sont présentés dans le Tableau 24.

<b>Paramètre</b>	<b>Corrélation avec la somme des anthocyanes moléculaires</b>	<b>R<sup>2</sup>P</b>
Dp-3-O-glc	0,86	0,68
Cy-3-O-glc	0,35	0,09
Pt-3-O-glc	0,90	0,76
Pn-3-O-glc	0,60	0,27
Mlv-3-O-glc	0,98	0,86
Dp-3-O-acglc	0,84	0,64
Cy-3-O-acglc	0,57	0,32
Pt-3-O-acglc	0,90	0,72
Pn-3-O-acglc	0,81	0,52
Mlv-3-O-acglc	0,84	0,61
Pn-3-O-cmglc	0,90	0,72
Mlv-3-O-cmglc	0,91	0,78

Tableau 24 : Corrélation des différentes anthocyanes avec la somme des anthocyanes moléculaires et résultats de prédictions des différentes anthocyanes avec le modèle de prédiction de la somme des anthocyanes moléculaires (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, R<sup>2</sup>P : coefficient de régression de prédiction)

Nous observons une diminution de la prédiction en comparaison des modèles présentés dans le Tableau 23, cependant, de nombreux composés montrent des robustesses avec la prédiction utilisant le modèle de la somme des anthocyanes qui se rapprochent de leur modèle respectif. Nous observons par exemple pour la malvidine-3-O-glucoside un R<sup>2</sup>CV de 0,88 pour sa modélisation (Tableau 23), et un R<sup>2</sup>P de 0,86 quand celle-ci est prédite par le modèle des anthocyanes totales (Tableau 24). Nous pouvons donc affirmer que pour ce composé, sa prédiction est due à sa corrélation avec les anthocyanes totales, et le modèle n'est donc pas représentatif de la concentration de la molécule recherchée. Cette observation peut s'étendre à toutes les anthocyanes moléculaires, car nous remarquons un lien important entre la robustesse de chaque modèle et sa corrélation avec la somme des anthocyanes totales. Ainsi, la concentration de l'anthocyane la moins bien modélisée, la cyanidine-3-O-acétylglucoside (R<sup>2</sup>CV de 0,58) montre une corrélation avec la concentration des

anthocyanes totales faible (0,57). Tout comme pour le suivi de la maturité du raisin, nous ne pouvons pas considérer les modèles de chaque anthocyane moléculaire comme fiables et construits sur la réponse spectrale du composé recherché. Ces résultats ne signifient pas que ces modèles ne sont pas exploitables, mais que ceux-ci doivent être soumis à la discussion car ils sont basés sur une prédiction indirecte. Plusieurs études ont montré des résultats pour la prédiction des anthocyanes similaires sans discuter de la pertinence de la robustesse (13,287). L'étude de notre jeu d'échantillons pendant la maturité du raisin et la fermentation démontre que même si les modèles de prédiction créés sont en apparence robustes, il est nécessaire de rechercher les possibles corrélations entre composés pour s'affranchir d'une surinterprétation des résultats.

Nous avons pu suivre l'évolution de la polymérisation des anthocyanes grâce aux valeurs des pigments monomériques (MP), petits pigments polymérisés (SPP) et grands pigments polymérisés (LPP). Les résultats de prédiction obtenus pour les MP montrent une bonne robustesse, avec un  $R^2CV$  de 0,86 et un RPD CV de 2,71 (Tableau 23). Cependant, la prédiction des SPP et LPP est discutable, avec un  $R^2CV$  respectivement de 0,74 et 0,68, et un RPD CV de 1,95 et 1,79. En se référant au suivi de ces paramètres pendant la fermentation (Figure 66), nous remarquons que les proportions de SPP et LPP sont bien plus faibles que celles de MP, ce qui peut expliquer cette perte de robustesse. De plus, la distinction faite par précipitation à l'albumine de sérum bovin entre les deux types de polymères d'anthocyanes n'est pas forcément discernable en IRTF. Intéressons-nous alors aux pigments polymérisés (PP, somme de SPP et LPP). Le modèle montre une prédiction supérieure aux SPP et LPP, avec un  $R^2CV$  de 0,83 et un RPD CV de 2,36. Les deux modèles MP et PP montrent des résultats assez robustes pour être utilisés, et leur utilisation combinée permet d'obtenir le pourcentage de pigments monomériques et de pigments polymérisés dans l'échantillon, et ainsi de suivre la polymérisation des anthocyanes pendant la fermentation. Concernant l'évolution des anthocyanes, le dernier paramètre modélisé est la concentration en anthocyanes libres, obtenue par décoloration au bisulfite de sodium. Les résultats sont très satisfaisants, avec un  $R^2CV$  de 0,89 et un RPD CV de 2,80. L'erreur de validation croisée (RMSECV), qui reflète fortement l'erreur de prédiction, est de 48,25 mg/L et apparaît tout à fait acceptable en comparaison de la moyenne obtenue sur le jeu d'échantillons (448,15 mg/L). Ces résultats peuvent être comparés aux travaux d'Aleixandre-Tudo et al., 2018, qui obtenait pour le même paramètre avec un modèle construit sur des vins en fermentation et finis un RMSECV de 56,5 mg/L. De plus, le nombre de variables latentes utilisées dans cette publication est de 24, contre 12 pour le modèle présenté dans le Tableau 23. S'il est difficile de comparer ces résultats

compte tenu de la différence entre les échantillons analysés, il apparaît que le modèle obtenu est plus performant que ceux présentés dans la littérature à ce jour.

Concernant la modélisation des tanins, nous constatons que pour les deux paramètres étudiés (précipitation à l'albumine de sérum bovin : BSA et précipitation à la méthylcellulose : MCP) nous observons les prédictions les plus robustes obtenues. Seul le résultat de MCP est comparable avec la littérature avec un travail similaire comprenant une majorité de vins en fermentation (287). Le RMSECV avancé dans ce travail est de 224 mg/L avec 23 variables latentes contre 178,15 mg/L et 9 variables latentes pour la validation croisée présentée dans le Tableau 23, ce qui démontre une performance accrue de notre modèle. Les résultats obtenus pour BSA apparaissent légèrement supérieurs à ceux de MCP, avec un RPD CV de 3,90 pour BSA et 3,65 pour MCP. Ces modèles mettent en évidence le fort potentiel de l'IRTF pour prédire les tanins pendant la fermentation, et doivent être désormais confrontés à une validation par millésime afin de s'assurer qu'ils resteront fiables d'une année à l'autre.

*(2) Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification par UV-visible*

Le Tableau 25 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres UV-visible.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL
<b>Composition en flavan-3-ols</b>						
Flavan-3-ol (mmol/L)	0,85	2,66	0,83	2,8	2,42	10
Flavan-3-ol terminal (mmol/L)	0,68	1,26	0,64	1,32	1,62	9
Catéchine (mmol/L)	0,81	0,58	0,78	0,63	2,10	10
Epicatéchine (mmol/L)	0,86	1,41	0,85	1,46	2,56	7
Epigallocatechine (mmol/L)	0,63	1,11	0,61	1,14	1,53	10
Epicatéchine-gallate (mmol/L)	0,79	0,16	0,77	0,17	2,12	7
<b>Anthocyanes moléculaires</b>						
Dp-3-O-glc (mg/L)	0,74	6,67	0,70	7,15	1,81	9
Cy-3-O-glc (mg/L)	0,55	1,94	0,50	2,05	1,40	9
Pt-3-O-glc (mg/L)	0,75	6,25	0,69	6,92	1,78	9
Pn-3-O-glc (mg/L)	0,41	6,89	0,38	7,08	1,21	7
Mlv-3-O-glc (mg/L)	0,77	23,66	0,74	25,25	1,94	7
Dp-3-O-acglc (mg/L)	0,58	2,22	0,53	2,33	1,45	7
Cy-3-O-acglc (mg/L)	0,61	0,79	0,57	0,82	1,43	11
Pt-3-O-acglc (mg/L)	0,73	1,34	0,71	1,38	1,82	8
Pn-3-O-acglc (mg/L)	0,62	1,86	0,59	1,93	1,53	7
Mlv-3-O-acglc (mg/L)	0,71	8,79	0,67	9,34	1,72	6
Pn-3-O-cmglc (mg/L)	0,70	1,23	0,68	1,28	1,75	8
Mlv-3-O-cmglc (mg/L)	0,76	4,06	0,72	4,36	1,87	8
Total glucoside (mg/L)	0,74	39,95	0,73	40,85	1,90	7
Total acétyl (mg/L)	0,75	12,03	0,72	12,72	1,86	7
Total coumaroyl (mg/L)	0,77	4,86	0,74	5,21	1,93	8
Total anthocyanes (mg/L)	0,76	53,65	0,74	55,29	1,95	7
<b>Evolution des anthocyanes</b>						
MP (Absorbance)	0,90	0,20	0,88	0,20	3,21	8
SPP (Absorbance)	0,64	0,14	0,62	0,14	1,63	7
LPP (Absorbance)	0,54	0,11	0,51	0,11	1,48	5
PP (Absorbance)	0,77	0,16	0,73	0,16	2,09	6
Anthocyanes libres (mg/L)	0,89	45,97	0,87	49,29	2,74	8
<b>Tanins</b>						
BSA (mg/L)	0,90	72,67	0,89	73,17	3,15	2
MCP (mg/L)	0,93	174,24	0,92	179,47	3,62	8

Tableau 25 : Modèles de prédiction PLS par UV-visible avec validation croisée pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation des raisins (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, , VL : nombre de variables latentes)

En comparant ces résultats avec ceux obtenus pour l'IRTF Tableau 23, nous constatons que seul le modèle pour les pigments monomériques (MP) apparaît plus performant pour l'UV-visible. La

tendance observée pour l'analyse de la maturité du raisin se confirme et l'IRTF présente de meilleurs résultats que l'UV-visible dans la prédiction des anthocyanes et des tanins du raisin, moût et vin. Bien que moins robuste, la prédiction utilisant l'UV-visible est toutefois efficace et de nombreux paramètres présentent des RPD CV supérieurs à 2.

Concernant la composition en flavan-3-ols, nous observons de bonnes performances pour la concentration en unités flavan-3-ols totales, avec un  $R^2CV$  de 0,83 et un RPD CV de 2,42. Le manque de précision obtenue par la concentration en unités flavan-3-ols terminales ( $R^2CV$  de 0,64 et un RPD CV de 1,62) ne permet pas, tout comme l'IRTF, de prédire de manière fiable le degré de polymérisation moyen. Pour les concentrations de chaque unité flavan-3-ol, nous observons de bons résultats pour l'épicatéchine ( $R^2CV$  de 0,85 et un RPD CV de 2,56), et des résultats corrects pour la catéchine ( $R^2CV$  de 0,78 et un RPD CV de 2,10) et l'épicatéchine-gallate ( $R^2CV$  de 0,77 et un RPD CV de 2,12). Cependant, pour l'épicatéchine-gallate, ces résultats plus faibles que l'IRTF peuvent être sujets à discussion car ils manquent de robustesse pour espérer prédire les proportions de tanins de pépins présentes dans l'échantillon.

Les résultats des concentrations en anthocyanes moléculaires montrent le même schéma que ceux obtenus pour l'IRTF, c'est-à-dire que les anthocyanes les mieux prédites sont celles fortement corrélées avec la somme des anthocyanes, et a contrario les anthocyanes qui disposent des modèles les moins prédictifs sont celles le moins bien corrélées avec la somme des anthocyanes. Nous ne considérerons donc pas les modèles obtenus pour chaque anthocyane comme fiables et nous ne discuterons que des anthocyanes totales, et de la somme des anthocyanes glucoside, acétylglucoside et coumaroylglucoside. Le modèle obtenu pour les anthocyanes totales est le plus performant, avec un  $R^2CV$  de 0,74 et un RPD CV de 1,95. Cependant, il n'apparaît pas suffisamment robuste pour être utilisé, avec une erreur de validation croisée de 55,29 mg/L que nous considérons trop importante en comparaison à la valeur moyenne de 303,64 mg/L obtenue pour le jeu d'échantillons. Le même constat peut être fait pour la somme des anthocyanes glucoside (RPD CV de 1,90), acétylglucoside (RPD CV de 1,86) et coumaroylglucoside (RPD CV de 1,93). Si ces modélisations peuvent toujours être utilisées comme des indicateurs, on leur préférera le modèle des anthocyanes libres ( $R^2CV$  de 0,87 et un RPD CV de 2,74) qui montre une meilleure robustesse. Pour expliquer cette différence, l'hypothèse est que l'UV-visible peut être influencé plus fortement par des facteurs comme la copigmentation. La valeur d'anthocyanes libres qui représente la couleur induite par les anthocyanes non polymérisées sera donc plus facile à modéliser que la valeur d'anthocyanes moléculaires, plus spécifique à la composition moléculaire.

Pour l'évolution des anthocyanes, tout comme l'IRTF, seuls les modèles des pigments monomériques (MP) et polymériques (PP) montrent une prédiction suffisante pour être utilisés. Le modèle des MP apparaît particulièrement robuste, avec un  $R^2CV$  de 0,88 et un RPD CV de 3,21, supérieur à l'IRTF. Le modèle des PP est moins performant ( $R^2CV$  de 0,73 et un RPD CV de 2,09), et sa prédiction devra être confirmée par la validation par millésime, afin de s'assurer de pouvoir être utilisée pour un suivi efficace de la polymérisation des anthocyanes durant la fermentation.

Concernant la prédiction de la concentration en tanins, nous observons de nouveau de très bons résultats. Contrairement à l'IRTF, ce sont ici les tanins précipitables à la méthylcellulose (MCP) qui apparaissent mieux prédits que les tanins précipitables à l'albumine de sérum bovin (BSA). Le modèle pour MCP indique une valeur de  $R^2CV$  de 0,92 et un RPD CV de 3,62, alors que le modèle pour BSA indique  $R^2CV$  de 0,89 et un RPD CV de 3,15.

De manière générale, l'UV-visible montre des prédictions légèrement inférieures à celles obtenues par IRTF, avec cependant des résultats très satisfaisants pour un bon nombre de paramètres, tels que la concentration en unités flavan-3-ols, en catéchine, épicatechine et épicatechine-gallate, la teneur en pigments monomériques et polymériques, la concentration en anthocyanes libres ainsi que la concentration en tanins précipitables à la méthylcellulose et à l'albumine de sérum bovin. Ces résultats doivent maintenant être confirmés par la validation par millésime, mais il est intéressant de regarder avant cela si la combinaison de l'IRTF et l'UV-visible peut améliorer la prédiction de certains paramètres.

### *(3) Construction de modèles de prédiction PLS avec validation croisée pour le suivi des anthocyanes et des tanins pendant la vinification par combinaison de l'IRTF et l'UV-visible*

Le Tableau 26 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres IRTF et UV-visible. Considérant que les anthocyanes moléculaires souffrent d'une corrélation indirecte avec les anthocyanes totales, comme nous l'avons démontré avec les spectres infrarouges, nous ne présenterons que la modélisation des anthocyanes totales, et de la somme des anthocyanes glucoside, acétylglucoside et coumaroylglucoside.

Paramètre	R <sup>2</sup> C	RMSEC	R <sup>2</sup> CV	RMSECV	RPD CV	VL
<b>Composition en flavan-3-ols</b>						
Flavan-3-ol (mmol/L)	0,86	2,51	0,85	2,68	2,53	10
Flavan-3-ol terminal (mmol/L)	0,75	1,12	0,70	1,21	1,77	11
Catéchine (mmol/L)	0,84	0,54	0,80	0,60	2,20	12
Epicatéchine (mmol/L)	0,89	1,25	0,88	1,31	2,85	10
Epigallocatechine (mmol/L)	0,71	0,98	0,68	1,04	1,67	12
Epicatéchine-gallate (mmol/L)	0,81	0,16	0,79	0,17	2,12	9
<b>Anthocyanes moléculaires</b>						
Total glucoside (mg/L)	0,77	37,44	0,75	39,19	1,98	7
Total acétyl (mg/L)	0,78	11,28	0,75	12,09	1,95	7
Total coumaroyl (mg/L)	0,79	4,63	0,77	4,91	2,05	7
Total anthocyanes (mg/L)	0,79	50,39	0,76	53,01	2,04	7
<b>Evolution des anthocyanes</b>						
MP (Absorbance)	0,90	0,20	0,89	0,22	3,01	9
SPP (Absorbance)	0,76	0,11	0,71	0,12	1,83	10
LPP (Absorbance)	0,58	0,11	0,51	0,12	1,44	7
PP (Absorbance)	0,80	0,15	0,74	0,17	1,96	7
Anthocyanes libres (mg/L)	0,92	39,41	0,89	45,79	2,95	9
<b>Tanins</b>						
BSA (mg/L)	0,93	58,39	0,92	63,94	3,61	8
MCP (mg/L)	0,93	161,38	0,92	173,94	3,74	10

Tableau 26 : Modèles de prédiction PLS par IRTF-UV-visible avec validation croisée pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Si l'on compare ces résultats avec ceux obtenus pour l'IRTF (Tableau 23), nous constatons deux paramètres qui apparaissent mieux prédits par la combinaison IRTF-UV-visible. Il s'agit des pigments monomériques et des tanins précipitables à la méthylcellulose (MCP). Concernant les pigments monomériques, nous obtenons un R<sup>2</sup>CV de 0,89 et RPD CV de 3,01 pour la combinaison, contre un R<sup>2</sup>CV de 0,86 et RPD CV de 2,71 pour l'IRTF seul. Cependant, si l'on s'intéresse au modèle obtenu pour l'UV-visible seul (RPD CV de 3,21), nous remarquons qu'il présente des résultats supérieurs à la combinaison des deux zones spectrales. L'addition de l'UV-visible à l'IRTF n'apparaît donc pas bénéficier à l'augmentation de la prédiction. La complémentarité des deux spectres ne semble pas être observée.

Concernant la prédiction de MCP, les résultats obtenus pour l'IRTF-UV-visible (R<sup>2</sup>CV de 0,92 et RPD CV de 3,74) sont supérieurs à ceux de l'IRTF seul (R<sup>2</sup>CV de 0,92 et RPD CV de 3,65) et de

l'UV-visible seul ( $R^2CV$  de 0,92 et un  $RPD CV$  de 3,62). Cependant, ils n'apparaissent pas suffisamment élevés pour considérer l'utilisation des deux méthodes spectroscopiques comme viable face à l'utilisation d'un seul appareillage IRTF ou UV-visible. De plus, nous constatons que ces données sont en contradiction avec celles obtenues sur le suivi de la maturité du raisin, où la combinaison IRTF-UV-visible donnait les résultats les plus faibles pour la prédiction de MCP.

Les modèles obtenus sur l'analyse de la fermentation ne nous permettent pas de conclure à une augmentation de la robustesse de prédiction pour la combinaison IRTF-UV-visible. Il apparaît que pour la majorité des modèles, l'utilisation de l'IRTF ou de l'UV-visible seul est plus efficace. De plus, pour les deux paramètres montrant une augmentation de prédiction, celle-ci n'est pas assez élevée pour justifier l'utilisation des deux méthodes spectroscopiques. Au regard de ces résultats, nous n'investiguerons pas la combinaison IRTF-UV-visible pour la validation des modèles en fonction des millésimes pour le suivi de la fermentation.

### ***b) Construction de modèles de prédiction PLS avec validation externe pour le suivi des anthocyanes et des tanins pendant la vinification des moûts de raisin***

Nous avons pu constater lors de l'analyse des échantillons de la fermentation que le millésime avait un impact sur la distribution du jeu de données. De même, lors de l'analyse du raisin pendant la maturité l'effet millésime est important et peut agir sur la robustesse des modèles. Pour valider nos modélisations il est donc nécessaire de prendre en compte cette variabilité. Les modèles de prédiction ont été construits avec deux des millésimes étudiés, et nous avons utilisé le dernier millésime à notre disposition comme jeu de données de validation externe.

#### ***(1) Construction de modèles de prédiction PLS avec validation par millésime pour le suivi des anthocyanes et des tanins pendant la vinification des moûts de raisin par IRTF***

Le Tableau 27 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres IRTF avec validation par millésime. Pour la validation 2017, les millésimes 2018 et 2019 ont été utilisés pour la calibration. Pour la validation 2018, ce sont ceux de 2017 et 2019 qui furent utilisés et pour la validation 2019 ceux de 2017 et 2018.

Dans un premier temps la comparaison des résultats obtenus pour chaque millésime a été étudiée. Nous constatons une forte disparité des prédictions en fonction des années utilisées pour la calibration

et la validation. Pour la composition en flavan-3-ols, les données pour la validation 2018 apparaissent inférieures à celles de 2017 et 2019. Rappelons qu'un modèle avec un RPD inférieur à 1,4 ne peut pas être considéré comme fiable, il faut que celui-ci soit au minimum compris entre 1,4 et 2 pour être envisagé pour la prédiction de la composition étudiée (10). Les résultats pour les modèles de prédiction en unités flavan-3-ols totales (Flavan-3-ol), indiquent des résultats concluants et exploitables pour la validation 2017 (RPD P : 1,98) et pour la validation 2019 (RPD P : 1,96) mais ils manquent de robustesse pour la validation 2018 (RPD P : 1,33). Nous observons la même tendance pour la concentration en unités flavan-3-ols terminales, en catéchine, épicatechine, épigallocatechine et épicatechine-gallate. Cependant, trois de ces paramètres montrent un RPD P supérieur à 1,4 pour toutes les validations : la concentration en catéchine (RPD P 2017 : 1,74 ; RPD P 2018 : 1,44 ; RPD P 2019 : 1,82), en épicatechine (RPD P 2017 : 2,02 ; RPD P 2018 : 1,95 ; RPD P 2019 : 2,81) et épicatechine-gallate (RPD P 2017 : 1,79 ; RPD P 2018 : 1,48 ; RPD P 2019 : 1,66). Si ces résultats apparaissent légèrement insuffisants pour être utilisés en routine dans les caves, ils montrent néanmoins la possibilité de prédire des concentrations (unités flavan-3-ols totales, catéchine, épicatechine et épicatechine-gallate) et ce malgré un effet millésime important. Pour aboutir à des modèles avec une erreur acceptable, il serait nécessaire de compléter notre base de données avec d'autres millésimes. Cela permettrait d'ajouter une variabilité supplémentaire et d'affiner la prédiction.

Paramètre	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL
	<b>Validation 2017</b>				<b>Validation 2018</b>				<b>Validation 2019</b>			
<b>Composition en flavan-3-ols</b>												
Flavan-3-ol (mmol/L)	0,76	3,17	1,98	10	0,70	5,27	1,33	11	0,87	2,96	1,96	8
Flavan-3-ol terminal (mmol/L)	0,63	1,18	1,60	12	0,32	2,89	0,85	11	0,71	1,69	0,89	8
Catéchine (mmol/L)	0,69	0,75	1,74	8	0,76	0,96	1,44	7	0,85	0,61	1,82	8
Epicatéchine (mmol/L)	0,77	1,82	2,02	10	0,84	1,92	1,95	7	0,89	1,21	2,81	7
Epigallocatechine (mmol/L)	0,61	0,71	1,49	10	0,21	3,02	0,70	7	0,53	1,48	0,74	8
Epicatéchine-gallate (mmol/L)	0,71	0,21	1,79	7	0,77	0,24	1,48	8	0,74	0,18	1,66	8
<b>Anthocyanes moléculaires</b>												
Total glucoside (mg/L)	0,69	51,60	1,34	11	0,32	54,87	1,10	5	0,64	48,70	1,36	11
Total acétyl (mg/L)	0,72	13,36	1,57	7	0,53	14,02	1,40	5	0,59	16,04	1,54	9
Total coumaroyl (mg/L)	0,73	6,96	1,19	13	0,69	5,69	1,74	13	0,51	7,10	1,41	8
Total anthocyanes (mg/L)	0,72	69,98	1,33	11	0,40	72,77	1,20	5	0,70	64,10	1,54	12
<b>Evolution des anthocyanes</b>												
MP (Absorbance)	0,83	0,28	2,38	6	0,82	0,33	2,31	3	0,70	0,34	1,68	4
SPP (Absorbance)	0,61	0,15	1,55	5	0,63	0,15	1,59	8	0,56	0,13	1,42	4
LPP (Absorbance)	0,57	0,13	1,29	8	0,50	0,12	1,34	10	0,51	0,12	1,36	8
PP (Absorbance)	0,65	0,22	1,65	7	0,76	0,19	1,91	9	0,62	0,19	1,48	8
Anthocyanes libres (mg/L)	0,74	85,40	1,72	11	0,73	65,89	1,88	14	0,71	73,28	1,73	7
<b>Tanins</b>												
BSA (mg/L)	0,90	125,48	2,07	8	0,94	64,16	3,45	11	0,93	71,77	2,76	7
MCP (mg/L)	0,88	253,24	2,78	8	0,88	257,43	2,46	7	0,91	180,51	3,16	7

Tableau 27 : Modèles de prédiction PLS par IRTF avec validation par millésimes pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Concernant la prédiction des anthocyanes moléculaires, nous observons des résultats de prédictions très inférieurs à ceux de la validation croisée présentée précédemment. La prédiction des anthocyanes apparaît donc très fortement impactée par l'effet millésime. Très peu des modèles construits dépassent le seuil du RPD P supérieur à un 1,4. Pour la somme des anthocyanes glucosides (Total glucoside), aucun des modèles obtenu ne montre une robustesse suffisante (RPD P 2017 : 1,34 ; RPD P 2018 : 1,10 ; RPD P 2019 : 1,36). Pour la somme des anthocyanes coumaroyls (Total coumaroyl), les prédictions obtenues apparaissent supérieures, avec seulement le modèle pour la validation 2017 qui ne montre pas une fiabilité suffisante (RPD P 2017 : 1,19 ; RPD P 2018 : 1,74 ; RPD P 2019 : 1,41). Pour les anthocyanes acétyls, tous les modèles peuvent être considérés comme prédictifs (RPD P 2017 : 1,57 ; RPD P 2018 : 1,40 ; RPD P 2019 : 1,54). Étonnamment, les modèles de la somme des anthocyanes moléculaires montrent des prédictions assez faibles (RPD P 2017 : 1,33 ; RPD P 2018 : 1,20 ; RPD P 2019 : 1,54), inférieures à celles obtenues pour les anthocyanes acétyls et coumaroyls. Il est possible d'envisager qu'avec le fort effet millésime constaté, les groupements acétyls ou coumaroyls caractéristiques de certaines anthocyanes soient un marqueur du spectre IRTF, ce qui expliquerait pourquoi celles-ci sont mieux prédites. Cependant, cette hypothèse ne peut être vérifiée avec le jeu de données actuel. Il serait intéressant d'ajouter à cette étude des vins en fermentation issus d'autres cépages et modalités de vinification avec des taux en anthocyanes glucosides, acétyls et coumaroyls différents de ceux déjà étudiés, ou alors de procéder à des ajouts dosés d'anthocyanes pour confirmer cette hypothèse.

Si l'on s'intéresse à l'évolution des anthocyanes, nous remarquons que plusieurs résultats sont exploitables. En effet, les prédictions des pigments monomériques (MP) indiquent une bonne robustesse, avec une légère baisse de précision pour le millésime 2019 (RPD P 2017 : 2,38 ; RPD P 2018 : 2,31 ; RPD P 2019 : 1,68). Ces résultats peuvent être couplés à ceux des pigments polymérisés (PP) qui montrent également un RPD P supérieur à 1,4 pour toutes les validations étudiées (RPD P 2017 : 1,65 ; RPD P 2018 : 1,91 ; RPD P 2019 : 1,48). L'utilisation de ces deux paramètres combinés permet de suivre l'évolution de la polymérisation des anthocyanes. Concernant la discrimination entre petits pigments polymérisés (SPP) et grands pigments polymérisés (LPP), nous observons des résultats divergents. Les SPP montrent des résultats inférieurs aux PP mais restent fiables (RPD P 2017 : 1,55 ; RPD P 2018 : 1,59 ; RPD P 2019 : 1,42), contrairement aux LPP dont la prédiction n'est pas acceptable (RPD P 2017 : 1,29 ; RPD P 2018 : 1,34 ; RPD P 2019 : 1,36). Ces derniers sont fortement minoritaires, ce qui pourrait expliquer la difficulté à les prédire. De plus, comme les petits pigments polymérisés représentent la majorité des pigments polymérisés (Figure 66), il est possible

d'envisager que leurs prédictions soient dues à une corrélation avec l'ensemble des pigments polymérisés. Il est donc préférable de ne pas considérer le modèle des SPP comme fiable. Pour le suivi des anthocyanes libres par décoloration au bisulfite de sodium, nous constatons des résultats de prédictions par millésimes homogènes (RPD P 2017 : 1,72 ; RPD P 2018 : 1,88 ; RPD P 2019 : 1,73). Si ces résultats sont inférieurs à ceux obtenus pour la validation croisée (RPD CV : 2,80), la perte de prédiction apparaît être indépendante de l'année utilisée pour la validation. Ces résultats sont particulièrement intéressants car en ajoutant de nouvelles données à la calibration, nous devrions être en mesure d'obtenir un modèle de plus en plus robuste, à l'épreuve de l'effet millésime.

Les derniers paramètres étudiés sont ceux relatifs aux tanins totaux, mesurés grâce à la précipitation à l'albumine de sérum bovin (BSA) et à la méthylcellulose (MCP). Nous observons les meilleurs résultats de prédiction étudiés jusqu'à présent, avec un RPD P supérieur à 2 pour les deux paramètres pour toutes les validations par millésimes. Pour BSA, nous obtenons un RPD P 2017 de 2,07, un RPD P 2018 de 3,45 et un RPD P 2019 de 2,76. Pour MCP, nous obtenons un RPD P 2017 de 2,78, un RPD P 2018 de 2,46 et un RPD P 2019 de 3,16. Ces résultats nouveaux sont directement exploitables, et montrent que les modèles construits avec l'ensemble des données peuvent prédire efficacement un millésime externe à la calibration.

La construction de modèles de prédiction PLS avec validation par millésime nous a permis de mettre en évidence les limites de certaines prédictions présentées dans le Tableau 23, mais aussi de valider la prédiction de certains modèles. Ainsi, nous avons pu constater que les modèles créés pour les anthocyanes moléculaires souffrent d'un manque de robustesse, mais que les modèles pour la composition en flavan-3-ols et pour l'évolution des anthocyanes possèdent un réel potentiel de prédiction. De plus, les modèles créés pour le suivi des tanins pendant la fermentation sont d'ores et déjà exploitables avec une bonne prédiction ( $RPD P > 2$ ). Afin de finir d'évaluer le potentiel de ces modèles, nous avons décidé de construire des modèles de prédiction PLS avec validation externe déterminée par l'algorithme de Kennard-Stone, qui permet de diviser le jeu de données en un set de calibration et de validation de variabilité proche (340).

## *(2) Construction de modèles de prédiction PLS par IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone*

L'utilisation de la validation croisée permet de mettre en évidence les paramètres susceptibles d'être modélisés, et la validation externe par millésime nous a quant à elle permis d'étudier l'impact de l'année sur la robustesse de prédiction. Afin d'évaluer correctement le potentiel

de nos modèles, il apparaît nécessaire de les confronter à une validation externe qui reflète la variabilité de la calibration. Pour ce faire, nous avons utilisé l'algorithme de Kennard-Stone pour séparer le jeu de données en un set de calibration (66 % des échantillons) et un set de validation (34 % des échantillons). Cet algorithme sert à maximiser la cohérence entre les erreurs d'étalonnage et de validation, et permet d'obtenir un set de validation représentatif de la variabilité du jeu de données (340). Le Tableau 28 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone.

Paramètre	R <sup>2</sup> P	RMSEP	RPD P	VL
<b>Composition en flavan-3-ols</b>				
Flavan-3-ol (mmol/L)	0,73	2,76	2,46	8
Flavan-3-ol terminal (mmol/L)	0,60	1,19	1,80	8
Catéchine (mmol/L)	0,58	0,61	2,16	7
Epicatéchine (mmol/L)	0,75	1,43	2,62	8
Epigallocatechine (mmol/L)	0,72	0,90	1,93	11
Epicatéchine-gallate (mmol/L)	0,73	0,16	2,25	9
<b>Anthocyanes moléculaires</b>				
Total glucoside (mg/L)	0,79	27,49	2,82	13
Total acétyl (mg/L)	0,69	9,94	2,38	13
Total coumaroyl (mg/L)	0,68	3,96	2,54	13
Total anthocyanes (mg/L)	0,79	36,22	2,98	15
<b>Evolution des anthocyanes</b>				
MP (Absorbance)	0,54	0,30	2,14	8
PP (Absorbance)	0,62	0,14	2,48	16
Anthocyanes libres (mg/L)	0,46	55,07	2,45	14
<b>Tanins</b>				
BSA (mg/L)	0,88	58,47	3,94	8
MCP (mg/L)	0,86	231,76	2,80	8

Tableau 28 : modèles de prédiction PLS par IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Pour la validation par millésime, nous avons considéré les modèles de prédiction avec un RPD P supérieur à 1,40 comme fiables, car l'impact du millésime était fort et de tels résultats permettaient de montrer le caractère prédictif du modèle face à une forte variabilité. Avec cette validation externe déterminée par l'algorithme de Kennard-Stone, nous recherchons la performance du modèle, car le

jeu de validation montre une variabilité très proche du jeu de calibration. Ainsi, nous considérerons un modèle comme fiable avec un RPD P supérieur à 2.

Nous constatons que la majorité des modèles étudiés répondent à ce critère de fiabilité. Pour la composition en flavan-3-ol, seules les concentrations en unités flavan-3-ols terminales (RPD P : 1,80) et en épigallocatechine (RPD P : 1,93) n'apparaissent pas assez robustes. Les deux modèles qui montrent le plus d'intérêts pour le suivi de la vinification sont la concentration en unités flavan-3-ols totales (RPD P : 2,46) et la concentration en épicatechine-gallate (RPD P : 2,25). En effet, en combinant ces informations il est possible de suivre le ratio d'épicatechine-gallate par rapport au nombre d'unités flavan-3-ols totales, et donc d'estimer la proportion de tanins de pépins. Si les résultats proposés avec la validation par millésime ne permettent pas d'avancer une fiabilité suffisante pour utiliser directement les modèles, les résultats obtenus grâce à la validation externe déterminée par l'algorithme de Kennard-Stone montrent bien la possibilité de prédire ces paramètres. Ainsi, avec une base de données d'échantillons suffisante pour couvrir toute la variabilité de la fermentation, il sera possible de prédire l'extraction des flavan-3-ols et la proportion de tanins de pellicule et de pépins.

Pour le suivi de la concentration en anthocyanes moléculaires, nous constatons une hausse importante de la prédiction en comparaison de la validation par millésime. Nous obtenons ainsi un RPD P de 2,82 pour la concentration en anthocyanes glucosides, un RPD P de 2,38 pour la concentration en anthocyanes acétyls, un RPD P de 2,54 pour la concentration en anthocyanes coumaroyls et un RPD P de 2,98 pour la concentration en anthocyanes totales. Si les validations par millésime manquaient de robustesse, nous constatons qu'avec un modèle qui reflète la variabilité des échantillons prédits, les résultats obtenus sont bien supérieurs (Figure 70). Pour la concentration en anthocyanes totales, nous obtenons un modèle exploitable, dont l'erreur de prédiction (36,22 mg/L) est tout à fait acceptable pour obtenir une valeur rapide permettant d'estimer les anthocyanes présentes pendant la fermentation.

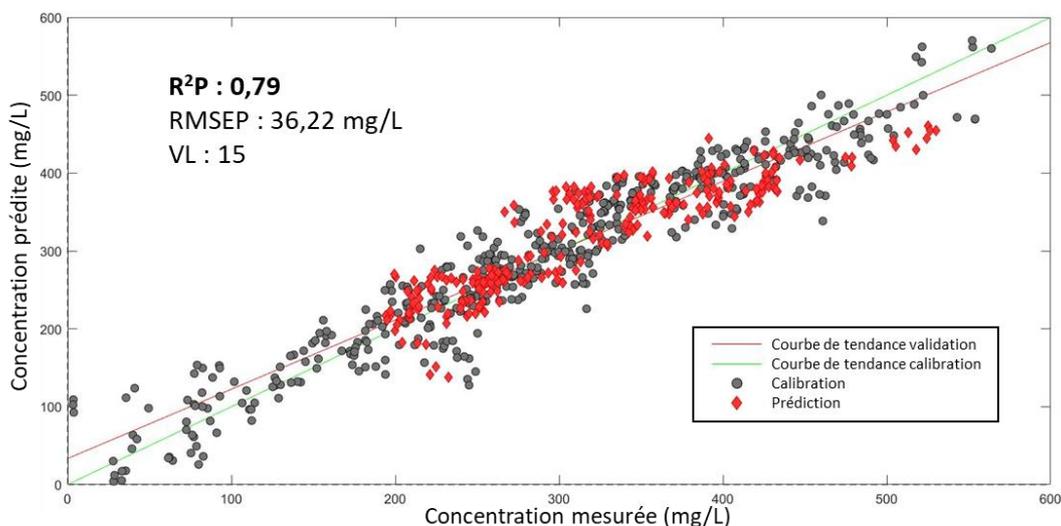


Figure 70 : modèle de prédiction de la somme des anthocyanes moléculaires avec validation externe déterminée par l’algorithme de Kennard-Stone sur les moûts de raisins en fermentation ( $R^2P$  : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, VL : nombre de variables latentes)

Sur la Figure 70 nous constatons que les échantillons prédits s’éloignent peu de la courbe de tendance de calibration, et nous n’observons pas d’échantillons déviants qui se détachent fortement du nuage de points. La somme des anthocyanes moléculaires est donc prédictible par IRTF. La validation par millésime indique cependant que la variabilité induite par l’année de vinification est forte, et ce modèle devra donc être complété avec d’autres échantillons afin de s’assurer de sa fiabilité.

Concernant l’évolution des anthocyanes, les résultats obtenus pour les pigments monomériques (RPD P : 2,14) sont très similaires à ceux obtenus pour la validation par millésimes (RPD P 2017 : 2,38 ; RPD P 2018 : 2,31 ; RPD P 2019 : 1,68). Des données aussi proches démontrent que la prédictibilité de ce paramètre est d’ores et déjà obtenue indépendamment du millésime. Pour les pigments polymériques, nous observons une hausse de la prédiction par rapport à la validation par millésime (RPD P 2017 : 1,65 ; RPD P 2018 : 1,91 ; RPD P 2019 : 1,48), avec un RPD P de 2,48, mais aussi une utilisation de plus de variables latentes (VL : 16) pour arriver à ce résultat. Le nombre de variables latentes utilisées pour la validation par millésime était de 7 pour 2017, 9 pour 2018 et 8 pour 2019. Cela signifie que même si la robustesse du modèle accroît avec un jeu de validation représentatif de la calibration, l’information utile extraite des spectres IRTF apparaît plus difficile à obtenir. Compte tenu de la variabilité observée pour l’évolution des anthocyanes (Figure 66), les modèles que nous avons obtenus apparaissent prometteurs, mais pas encore assez robustes pour être considérés comme fiables. Il sera nécessaire de les confronter à une variabilité accrue, notamment

avec des vins plus âgés où la concentration en pigments polymériques est plus importante, pour s'assurer de leur prédictivité.

Pour le modèle des anthocyanes libres, nous observons également une hausse de la prédiction avec la validation externe déterminée par l'algorithme de Kennard-Stone (RPD P : 2,45) par rapport à la validation par millésime (RPD P 2017 : 1,72 ; RPD P 2018 : 1,88 ; RPD P 2019 : 1,73). Ce modèle peut être directement comparé à celui de la somme des anthocyanes moléculaires, car ces deux paramètres mesurent des informations proches. Les anthocyanes libres mesurent la couleur induite par les anthocyanes non polymérisées, alors que la valeur d'anthocyanes moléculaires mesure la concentration en anthocyanes non polymérisées. Si ces deux paramètres peuvent paraître redondants, nous avons pu préciser précédemment que certains facteurs tels que la copigmentation peuvent avoir un impact non négligeable sur la couleur sans pour autant influencer sur la concentration. Nous observons que la somme des anthocyanes moléculaires est mieux prédite (RPD P : 2,98) que les anthocyanes libres. L'analyse infrarouge se base sur l'absorbance spécifique des liaisons chimiques, et apparaît donc plus adaptée pour mesurer les teneurs moléculaires que l'impact colorimétrique engendré. A contrario, l'UV-visible mesure grâce à la partie visible les différences colorimétriques. Nous devrions donc observer une meilleure prédiction des anthocyanes libres par rapport à la somme des anthocyanes moléculaires. Cette hypothèse devra être vérifiée quand nous aborderons les résultats de l'UV-visible.

Les modèles de prédiction des tanins montrent de très bonnes performances, notamment pour la prédiction des tanins précipitables à l'albumine de sérum bovin (BSA), avec un RPD P de 3,94. Le modèle de prédiction des tanins dosés par précipitation à la méthylcellulose (MCP) montre une robustesse inférieure (RPD P : 2,80) mais s'avère néanmoins efficace. Ces données corroborent l'utilité de l'IRTF déjà observée lors de la validation par millésime pour prédire la concentration en tanins pendant la fermentation. La représentation graphique de la prédiction de ces deux modèles est présentée sur la Figure 71. Tout comme pour le modèle de la somme des anthocyanes moléculaires, nous n'observons pas d'échantillons fortement divergents de la courbe de tendance de calibration. Etant donné qu'aucun échantillon n'a été écarté, les modèles présentés fourniront donc une valeur en tanins prédits fiable (en prenant compte l'erreur de prédiction moyenne RMSEP), sans incohérences dans les concentrations obtenues. Nous observons toutefois une zone de saturation pour les deux modèles. En effet, lorsque les concentrations sont élevées (supérieures à 700 mg/L pour BSA et supérieures à 2000 mg/L pour MCP), les résultats de prédiction apparaissent très proches et s'éloignent de la concentration mesurée. Par rapport à l'ensemble des données, peu d'échantillons

appartiennent à cette zone de saturation. Nous observons un nombre d'échantillons plus important pour MCP que pour BSA, et cela pourrait expliquer pourquoi le modèle construit pour MCP montre une prédiction moins importante. Ces modèles nécessitent désormais d'être complétés avec des échantillons qui présentent des concentrations élevées en tanins, afin de vérifier si cette zone de saturation peut être atténuée avec de nouvelles données, ou si elle est induite par une saturation des spectres IRTF.

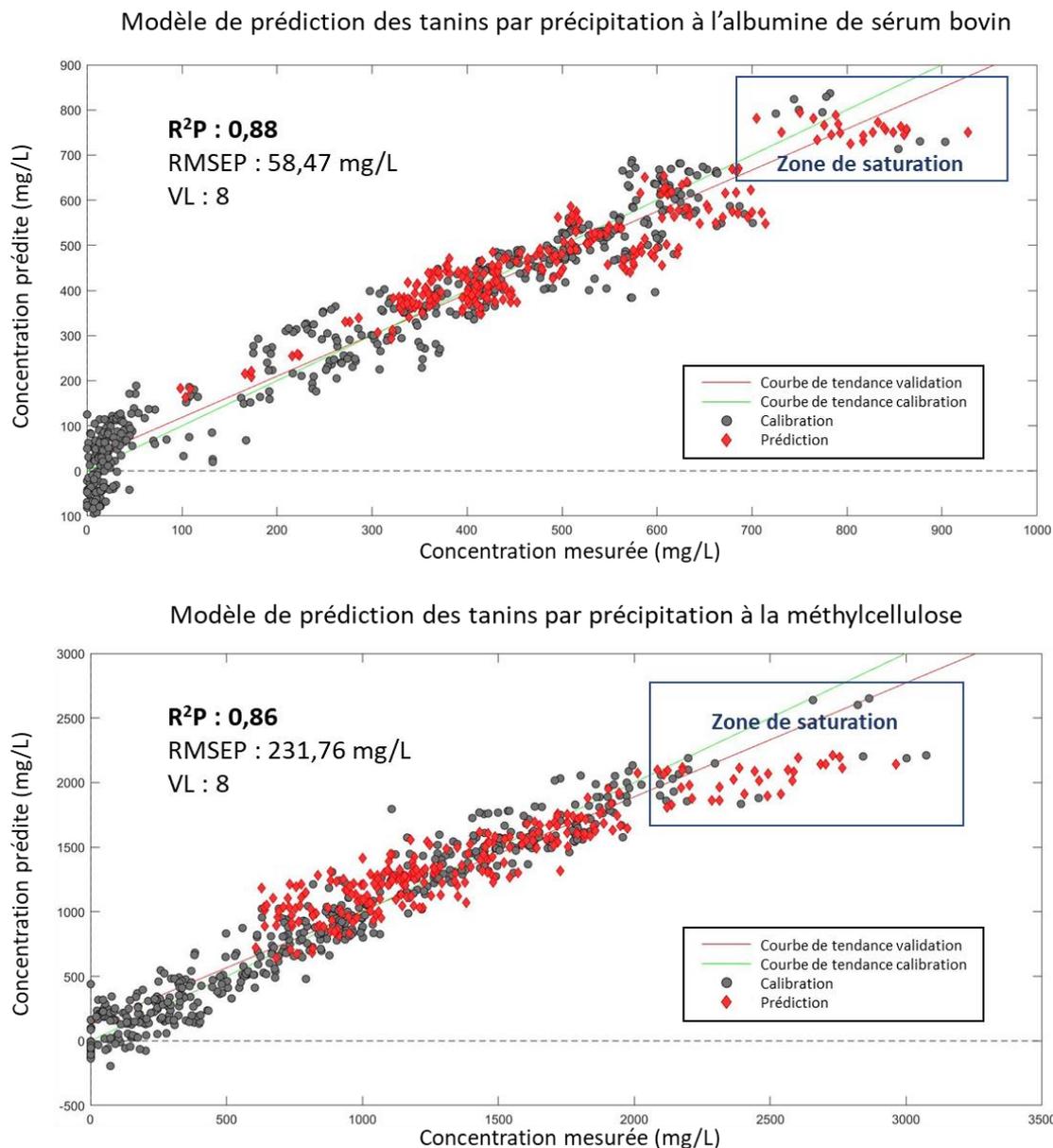


Figure 71 : modèles de prédiction des tanins précipitables à l'albumine de sérum bovin et la méthylcellulose avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation ( $R^2P$  : coefficient de régression de prédiction,  $RMSEP$  : erreur quadratique moyenne de prédiction,  $VL$  : nombre de variables latentes)

*(3) Construction de modèles de prédiction PLS avec validation par millésime pour le suivi des anthocyanes et des tanins pendant la vinification par UV-visible*

Le Tableau 29 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres UV-visible avec validation par millésime. Pour la validation 2017, les millésimes 2018 et 2019 ont été utilisés pour la calibration. Pour la validation 2018, ce sont ceux de 2017 et 2019 qui sont utilisés et pour la validation 2019 ceux de 2017 et 2018.

Paramètre	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL	R <sup>2</sup> P	RMSEP	RPD P	VL
	<b>Validation 2017</b>				<b>Validation 2018</b>				<b>Validation 2019</b>			
<b>Composition en flavan-3-ols</b>												
Flavan-3-ol (mmol/L)	0,19	9,56	0,66	8	0,72	5,59	1,26	10	0,71	3,15	1,84	11
Flavan-3-ol terminal (mmol/L)	0,01	6,23	0,30	8	0,31	3,06	0,80	10	0,61	1,77	0,85	11
Catéchine (mmol/L)	0,13	2,22	0,59	9	0,65	1,13	1,22	9	0,76	0,56	1,98	11
Epicatéchine (mmol/L)	0,60	2,89	1,27	8	0,83	1,92	1,95	10	0,80	1,49	2,28	8
Epigallocatechine (mmol/L)	0,06	3,60	0,29	9	0,32	2,75	0,77	9	0,52	1,40	0,78	5
Epicatéchine-gallate (mmol/L)	0,63	0,33	1,14	4	0,81	0,26	1,37	3	0,67	0,21	1,39	10
<b>Anthocyanes moléculaires</b>												
Total glucoside (mg/L)	0,33	116,90	0,59	16	0,51	55,24	1,09	6	0,57	69,53	0,96	7
Total acétyl (mg/L)	0,27	37,84	0,55	15	0,62	15,17	1,29	7	0,66	18,36	1,34	6
Total coumaroyl (mg/L)	0,32	17,45	0,47	16	0,57	6,86	1,44	7	0,55	8,43	1,19	6
Total anthocyanes (mg/L)	0,43	163,83	0,57	15	0,55	74,58	1,17	6	0,65	92,51	1,07	7
<b>Evolution des anthocyanes</b>												
MP (Absorbance)	0,79	0,37	1,76	7	0,87	0,28	2,64	9	0,64	0,34	1,66	8
SPP (Absorbance)	0,53	0,16	1,45	4	0,36	0,21	1,16	7	0,38	0,16	1,13	4
LPP (Absorbance)	0,52	0,12	1,38	4	0,58	0,12	1,36	4	0,42	0,15	1,11	4
PP (Absorbance)	0,42	0,27	1,33	4	0,70	0,23	1,52	7	0,55	0,23	1,21	4
Anthocyanes libres (mg/L)	0,80	74,52	1,97	4	0,79	58,82	2,11	11	0,80	63,24	2,00	8
<b>Tanins</b>												
BSA (mg/L)	0,88	100,64	2,59	2	0,95	58,56	3,78	7	0,90	64,66	3,06	3
MCP (mg/L)	0,93	244,73	2,88	3	0,87	238,62	2,66	6	0,89	245,05	2,33	8

Tableau 29 : Modèles de prédiction PLS par UV-visible avec validation par millésimes pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

En comparaison avec les résultats obtenus pour l'IRTF, les modèles construits autour de l'UV-visible apparaissent moins performants, notamment pour tous les paramètres inhérents à la composition moléculaire des polyphénols. Pour la composition en flavan-3-ols et en anthocyanes moléculaires, nous observons très peu de modèles avec un RPD P supérieur à 1,40. Le millésime semble fortement impacter la prédiction, avec des robustesses de modèle très divergentes d'une validation à l'autre. Si l'on s'intéresse par exemple aux modèles pour la concentration en unités flavan-3-ols totales, nous constatons un résultat prometteur pour la validation 2019 (RPD P : 1,84), mais celui-ci est contrebalancé par un modèle très peu performant pour la validation 2017 (RPD P : 0,66). Le même constat peut être fait pour les concentrations en unités catéchine (RPD P 2017 : 0,59 ; RPD P 2018 : 1,22 ; RPD P 2019 : 1,98) et épicatechine (RPD P 2017 : 1,27 ; RPD P 2018 : 1,95 ; RPD P 2019 : 2,28). Concernant les concentrations en unités épigallocatechine et épicatechine-gallate, les résultats obtenus sont inférieurs aux performances requises (RPD P > 1,40) pour s'avérer exploitables. La même observation est faite pour la composition en anthocyanes moléculaires, où seule la validation 2018 pour la somme des anthocyanes coumaroyls indique une possible prédictivité avec un RPD P de 1,44. L'étude bibliographique rapportait que les différents flavan-3-ols et anthocyanes possédaient des spectres d'absorption propres à chaque composé dans l'UV-visible (128,283). Cependant, au vu des résultats obtenus, il apparaît que la superposition d'information sur le spectre couplée à des variabilités différentes en fonction de l'année étudiée ne permettent pas d'extraire efficacement les informations recherchées.

Pour l'évolution des anthocyanes, nous observons de bons résultats pour les pigments monomériques (RPD P 2017 : 1,76 ; RPD P 2018 : 2,64 ; RPD P 2019 : 1,66). Cependant, pour les petits pigments polymérisés (RPD P 2017 : 1,45 ; RPD P 2018 : 1,16 ; RPD P 2019 : 1,13), les grands pigments polymérisés (RPD P 2017 : 1,38 ; RPD P 2018 : 1,36 ; RPD P 2019 : 1,11) et la somme des pigments polymérisés (RPD P 2017 : 1,33 ; RPD P 2018 : 2,52 ; RPD P 2019 : 1,21) les performances de prédiction sont inférieures, et n'apparaissent pas suffisantes pour être exploitées. L'intérêt de modéliser les teneurs en pigments monomériques seules est très limité, car il n'est alors pas possible de suivre de manière efficace la polymérisation des anthocyanes sans les teneurs en pigments polymériques.

Concernant la modélisation des anthocyanes libres, nous constatons des résultats directement exploitables, mais surtout une homogénéité de la prédiction indépendante du millésime (RPD P 2017 : 1,97 ; RPD P 2018 : 2,11 ; RPD P 2019 : 2,00). Ces données viennent valider l'hypothèse émise lors de la discussion des résultats l'IRTF. L'UV-visible apparaît tout à fait adapté à la prédiction des

anthocyanes libres mesurées par décoloration au bisulfite de sodium, qui mesure l'impact colorimétrique engendré par les anthocyanes non polymérisées, alors que l'IRTF est plus approprié pour mesurer les concentrations moléculaires.

Les modèles obtenus pour la prédiction des tanins montrent les meilleures performances observées pour la modélisation des polyphénols par les spectres UV-visible. Tout comme pour l'IRTF, Les meilleures prédictions sont obtenues pour les tanins précipitables à l'albumine de sérum bovin, avec un RPD P 2017 de 2,59, un RPD P 2018 de 3,78 et un RPD P 2019 de 3,06. Ces résultats apparaissent même supérieurs à l'IRTF, et montre la fiabilité de l'UV-visible dans la modélisation des tanins précipitables à l'albumine de sérum bovin. Pour les tanins précipitables à la méthylcellulose, la prédictibilité est légèrement inférieure (RPD P 2017 : 2,88 ; RPD P 2018 : 2,66 ; RPD P 2019 : 2,33), mais montre néanmoins une robustesse et une fiabilité suffisante pour être exploitable.

Pour s'affranchir de l'effet millésime et observer la prédictibilité des paramètres étudiés face à une validation externe qui reflète la variabilité de la calibration, nous avons, comme pour les modèles de prédiction PLS par IRTF, construit des modèles avec validation externe déterminée par l'algorithme de Kennard-Stone pour l'UV-visible.

*(4) Construction de modèles de prédiction PLS par IRTF avec validation externe déterminée par l'algorithme de Kennard-Stone*

Le Tableau 28 présente les résultats obtenus pour la modélisation des concentrations en anthocyanes et tanins grâce aux spectres UV-visible avec validation externe déterminée par l'algorithme de Kennard-Stone. Le jeu de données a été séparé en un set de calibration (66 % des échantillons) et un set de validation (34 % des échantillons).

Paramètre	R <sup>2</sup> P	RMSEP	RPD P	VL
<b>Composition en flavan-3-ols</b>				
Flavan-3-ol (mmol/L)	0,85	2,76	2,46	10
Flavan-3-ol terminal (mmol/L)	0,71	1,24	1,73	10
Catéchine (mmol/L)	0,83	0,57	2,32	11
Épicatéchine (mmol/L)	0,87	1,39	2,69	7
Épigallocatechine (mmol/L)	0,65	1,09	1,60	11
Épicatéchine-gallate (mmol/L)	0,79	0,16	2,25	10
<b>Anthocyanes moléculaires</b>				
Total glucoside (mg/L)	0,60	46,57	1,67	4
Total acétyl (mg/L)	0,62	13,80	1,71	7
Total coumaroyl (mg/L)	0,69	5,38	1,87	8
Total anthocyanes (mg/L)	0,62	62,77	1,72	7
<b>Evolution des anthocyanes</b>				
MP (Absorbance)	0,87	0,23	2,88	8
PP (Absorbance)	0,77	0,16	2,08	6
Anthocyanes libres (mg/L)	0,85	45,00	3,00	9
<b>Tanins</b>				
BSA (mg/L)	0,93	62,31	3,70	6
MCP (mg/L)	0,91	193,62	3,36	11

Tableau 30 : modèles de prédiction PLS par UV-visible avec validation externe déterminée par l'algorithme de Kennard-Stone pour les paramètres d'anthocyanes et de tanins suivis pendant la fermentation de moûts de raisins (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Pour la modélisation de la composition en flavan-3-ols, nous observons une hausse de la prédiction importante en comparaison à la validation par millésime, avec des résultats très proches de ceux obtenus pour l'IRTF. Nous obtenons ainsi plusieurs modèles avec un RPD P supérieur à 2, avec la concentration en unités flavan-3-ols totales (RPD P : 2,46), en catéchine (RPD P : 2,32), en épicatechine (RPD P : 2,69) et épicatechine-gallate (RPD P : 2,25). Cependant, contrairement à l'IRTF, les résultats obtenus lors de la validation par millésime pour l'UV-visible ne présageaient pas d'une prédiction efficace pour ces paramètres. Il semblerait donc que la variabilité du millésime influe très fortement sur les prédictions obtenues. La modélisation de ces composés par l'UV-visible peut donc être questionnée, car elle nécessite d'être construite avec une calibration composée d'un nombre important d'échantillons, couvrant toute la variabilité possible. De plus, cela n'assurerait pas l'obtention d'un modèle fiable. Au regard des résultats obtenus, l'IRTF apparaît bien plus adapté à la prédiction de la composition en flavan-3-ol pendant la fermentation.

Concernant les modèles de prédiction pour les anthocyanes moléculaires, nous constatons que malgré une hausse des résultats en comparaison à la validation par millésime, aucun modèle ne montre un RPD P supérieur à 2. Contrairement aux données de la littérature (12), les résultats que nous avons obtenus semblent montrer l'inaptitude de l'UV-visible à prédire de manière fiable les concentrations en anthocyanes moléculaires. Pour notre jeu de données, cette constatation est appuyée par le manque de robustesse observé sur les données de validation croisée, de validation externe par millésime et de validation externe déterminée par l'algorithme de Kennard-Stone.

Pour le suivi de l'évolution des anthocyanes, nous observons des résultats qui démontrent le fort potentiel de l'UV-visible. Tout d'abord, la prédiction des anthocyanes libres valide les résultats obtenus avec la validation par millésime, avec un RPD P de 3,00. Le modèle est présenté sur la Figure 72.

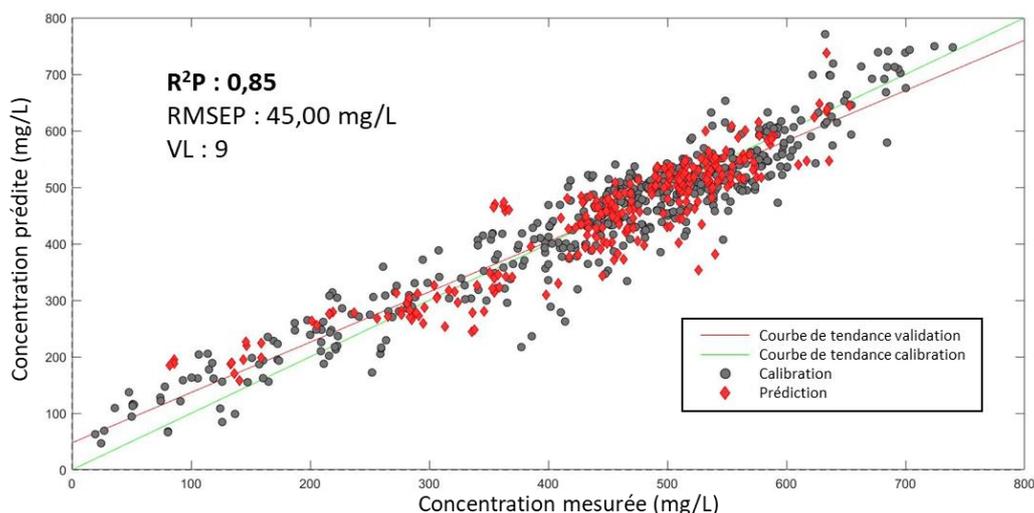


Figure 72 : modèle de prédiction des anthocyanes libres avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation ( $R^2P$  : Coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, VL : nombre de variables latentes)

Aucun échantillon fortement déviant n'est observable, et nous constatons que la distribution des échantillons de la validation s'intègre parfaitement au nuage de point de la calibration. Si l'on compare l'erreur de calibration obtenue pour le modèle construit avec tous les échantillons présentés dans le Tableau 25 (RMSECV : 45,97), avec l'erreur de prédiction du modèle de la Figure 72 (RMSEP : 45,00), nous remarquons qu'elles sont pratiquement identiques. Ces données confirment un modèle très fiable, qui ne souffre dans sa prédiction que de l'erreur induite par la calibration. Ainsi, si l'ajout de nouveaux échantillons permet d'obtenir une calibration plus performante, nous pourrions également augmenter la prédiction. Concernant les modèles des pigments monomériques et

polymériques, si la prédiction est moins efficace, nous obtenons toutefois une robustesse exploitable. Le modèle des pigments monomériques présente un RPD P de 2,88 et le modèle des pigments polymériques un RPD P de 2,08. Si le ratio des deux valeurs n'est pas suffisamment fiable pour le moment pour apporter une information non discutable, il peut d'ores et déjà servir d'indicateur pour estimer l'avancée de la polymérisation des anthocyanes. Ces modèles possèdent un réel potentiel, et doivent désormais être complétés avec d'autres échantillons pour augmenter leur fiabilité.

Enfin, les modèles des tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose sont les plus performants observés et sont présentés dans la Figure 73.

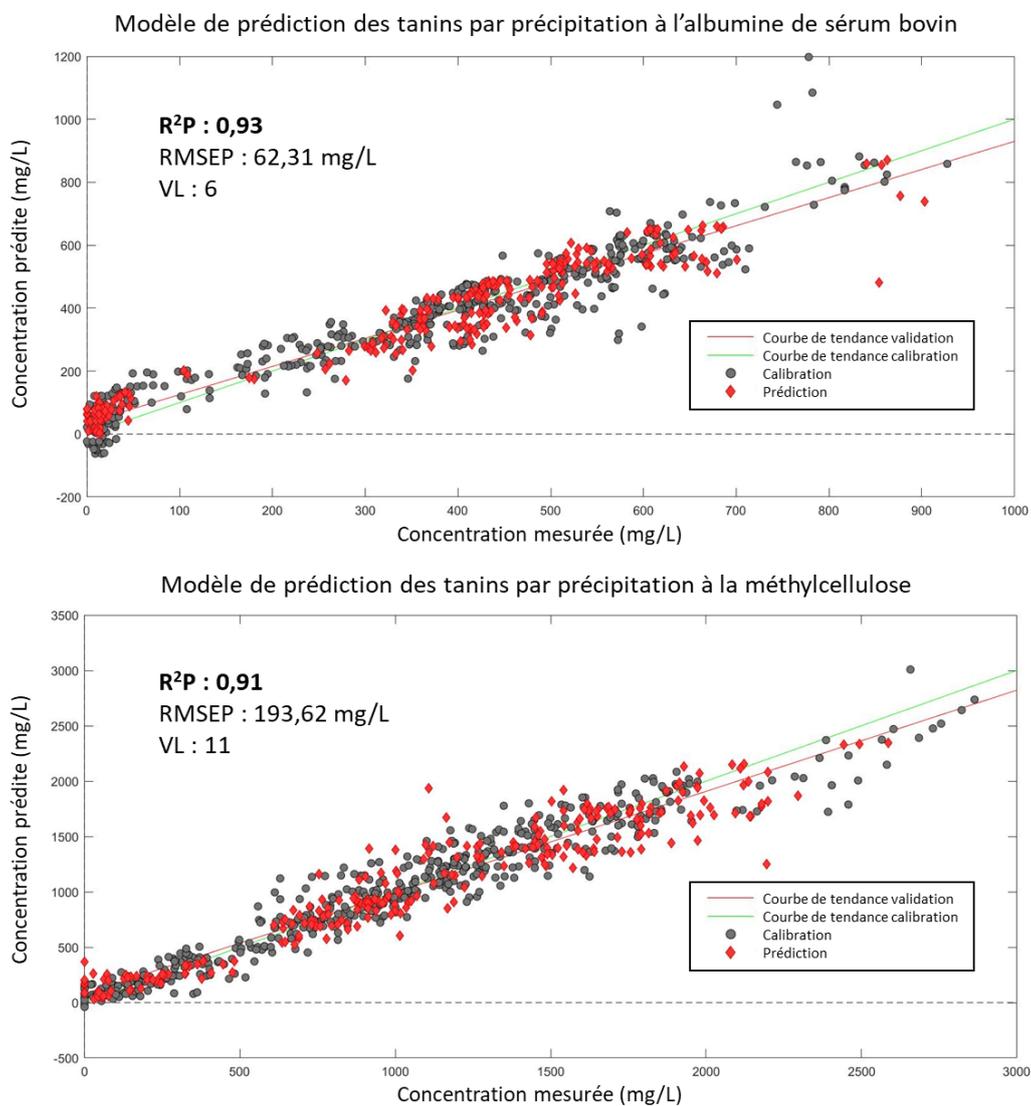


Figure 73 : modèles de prédiction des tanins précipitables à l'albumine de sérum bovin et la méthylcellulose avec validation externe déterminée par l'algorithme de Kennard-Stone sur les moûts de raisins en fermentation ( $R^2P$  : Coefficient de régression de prédiction,  $RMSEP$  : erreur quadratique moyenne de prédiction,  $VL$  : nombre de variables latentes)

Pour le modèle des tanins précipitables à l'albumine de sérum bovin, nous obtenons une excellente prédiction (RPD P : 3,70), mais nous constatons que comme pour l'IRTF, les concentrations hautes ne s'intègrent pas correctement à la linéarité des données. Contrairement à l'IRTF, qui montrait une zone de saturation, les échantillons de l'UV-visible s'écartent fortement de la courbe de tendance de calibration. Nous constatons également que cette divergence implique les échantillons de calibration et validation, et confirme donc la difficulté à prédire et modéliser les concentrations élevées. Cependant, aucune saturation n'a été constatée sur les spectres UV-visible, il pourrait donc s'agir d'un biais de manipulation ou de la complexité du spectre IRTF. Cependant, malgré ces échantillons déviants, le modèle propose une robustesse accrue, et valide donc l'utilisation de l'UV-visible pour prédire les tanins précipitables à l'albumine de sérum bovin. Le modèle de prédiction des tanins par précipitation à la méthylcellulose donne lui aussi d'excellents résultats, avec un RPD P de 3,36. Ces prédictions sont très proches de celles de la littérature obtenues pour des vins finis commerciaux (326), et pour un mélange de vins en fermentation et finis (12). Ces résultats apparaissent même supérieurs à ceux de l'IRTF, et nous n'observons pas de saturation pour les fortes concentrations. L'UV-visible est donc tout à fait adapté pour la prédiction des tanins, et les données obtenues via la validation par millésime indiquent que ce facteur n'influe pas de manière significative sur la fiabilité des modèles.

#### **4. Conclusions et perspectives de la conception de modèles de prédiction PLS pour l'analyse des anthocyanes et tanins du raisin pendant la fermentation**

Lors de cette étude portée sur la prédiction des tanins et des anthocyanes durant la fermentation par analyse PLS couplée à la spectroscopie IRTF et UV-visible, nous avons pu approfondir le potentiel et les limites de cette méthode. Nous avons tout d'abord étudié le jeu d'échantillons répartis sur 3 millésimes différents de vinification afin de s'assurer que la variabilité observée était optimale pour la conception de modèles de prédiction. Nous avons ainsi pu mettre en évidence la bonne diversité de l'échantillonnage, ainsi que comme pour l'étude de la maturité du raisin, un impact du millésime non négligeable sur les spectres IRTF et UV-visible. Les deux cépages étudiés, Merlot et Cabernet franc, ne montraient quant à eux pas de distinction notable qui pourrait influencer la modélisation.

Nous avons poursuivi en étudiant l'évolution caractéristique des différents paramètres de tanins et d'anthocyanes au cours de la fermentation. Dans l'ensemble, le jeu d'échantillons étudiés est apparu

tout à fait adapté à la conception de modèles de prédiction des tanins et des anthocyanes pendant la fermentation. Il montre une excellente variabilité, et l'étude des différents paramètres pendant la vinification indique des évolutions propres à chaque composé. Ainsi, compléter et construire des modèles de prédiction fiables permettrait d'obtenir des informations utiles au vinificateur pour contrôler la composition phénolique de ses vins pendant la fermentation.

Afin de s'assurer de la robustesse des modèles créés, nous avons décidé d'investiguer en profondeur la prédictibilité de chaque paramètre. Pour cela, nous avons tout d'abord développé des modèles avec tous les échantillons et une validation croisée. Cette première étape permet de mettre en évidence directement les paramètres exprimant le plus fort potentiel, et a contrario les paramètres qui seront difficiles à prédire. Ensuite, pour confronter les prédictions à l'impact du millésime, nous avons construit des modèles regroupant deux des millésimes étudiés, et utilisant le dernier millésime comme validation. Pour chaque paramètre, 3 modèles ont été construits afin d'utiliser chaque millésime étudié comme validation externe. Enfin, pour évaluer correctement le potentiel de nos modèles, nous les avons confrontés à une validation externe qui reflète la variabilité de la calibration, en divisant le jeu de données grâce à l'algorithme de Kennard-Stone.

Pour l'IRTF, les modèles avec validation croisée se sont avérés performants, notamment pour le suivi de la composition en flavan-3-ols, l'évolution des anthocyanes et les concentrations en tanins. Concernant les anthocyanes moléculaires, nous avons observé comme pour la maturité du raisin des résultats de prédiction qui apparaissaient extrêmement robustes par rapport à la faible concentration des composés étudiés. Nous avons donc recherché une possible surinterprétation des résultats et nous avons prouvé que les anthocyanes moléculaires étaient fortement corrélées avec la somme des anthocyanes, et que leurs prédictions découlaient de leur manque de variabilité et de cette corrélation. Nous avons ainsi confirmé les limites de l'analyse multivariée que nous avons abordées lors de la partie sur la maturité du raisin. Il est nécessaire de connaître le jeu de données étudié, et de bien comprendre le fonctionnement de l'analyse multivariée pour s'assurer de l'obtention de modèles de prédiction efficaces et fiables, sans mauvaise interprétation des résultats obtenus.

Concernant l'UV-visible, nous avons obtenu avec la validation croisée de bons résultats pour la composition en flavan-3-ols, l'évolution des anthocyanes et les concentrations en tanins. Pour les anthocyanes moléculaires les prédictions obtenues ne se sont pas avérées suffisantes pour être envisagées.

Nous avons aussi investigué l'effet de la combinaison de l'IRTF et l'UV-visible pour essayer de produire des modèles plus performants. Les résultats obtenus ne nous permettent pas de conclure à une augmentation de la robustesse de prédiction pour la combinaison IRTF-UV-visible. Il apparaît que pour la majorité des modèles, l'utilisation de l'IRTF ou de l'UV-visible seul est plus efficace.

Pour la validation par millésime, la différence de variabilité en fonction de l'année étudiée impacte fortement la prédictibilité des paramètres étudiés. Avec l'IRTF, nous avons toutefois observé des modèles avec un fort potentiel pour la prédiction de la concentration en unités flavan-3-ols totales, en catéchine, en épicatechine-gallate en anthocyanes libres, et pour la composition en pigments monomériques et polymérisés. De plus, les résultats obtenus pour les tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose montrent de très bonnes performances et sont directement exploitables. Les données observées pour les anthocyanes moléculaires approchent du seuil d'acceptabilité, et méritent d'être approfondies.

Les résultats obtenus avec l'UV-visible apparaissent moins performants que ceux de l'IRTF. Les modèles conçus pour la composition en flavan-3-ols et pour les anthocyanes moléculaires montrent une efficacité insuffisante. Pour les pigments monomériques nous obtenons de bonnes robustesses de validation, et des résultats encourageants pour les pigments polymériques. Nous avons obtenu des modèles assez prédictifs pour être directement exploités pour la concentration en anthocyanes libres et en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose.

Enfin, nous avons utilisé la modélisation avec validation externe déterminée par l'algorithme de Kennard-Stone pour identifier les modèles possédant un fort potentiel si la calibration s'affranchit de la variabilité induite par le millésime. Ces modèles pourront être développés mais nécessitent l'ajout de nouveaux échantillons pour affiner la calibration. Pour l'IRTF, nous constatons de bons résultats pour la majorité des paramètres. Nous obtenons un RPD P supérieur à 2 pour la concentration en unités flavan-3-ols, catéchine, épicatechine, épicatechine-gallate, pour la concentration de la somme des anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales, ainsi que pour la composition en pigments monomériques et polymériques, en anthocyanes libres et pour les tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose. Pour l'UV-visible, de bons résultats sont obtenus pour la concentration en unités flavan-3-ols, catéchine, épicatechine, épicatechine-gallate, et pour la composition en pigments monomériques et polymériques, en anthocyanes libres et pour les tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose.

De manière générale, l'IRTF apparaît être tout à fait adapté à la prédiction des anthocyanes et des tanins pendant la fermentation, et trouve un intérêt accru pour le détail de la composition moléculaire. Les modèles construits avec l'UV-visible apparaissent légèrement moins performants, mais sont particulièrement efficaces pour modéliser les informations généralistes, telles que la concentration en anthocyanes libres ou la concentration en tanins. Un récapitulatif présente les limites et les intérêts des deux méthodes spectroscopiques pour la prédiction des anthocyanes et des tanins pendant la fermentation (Figure 74).

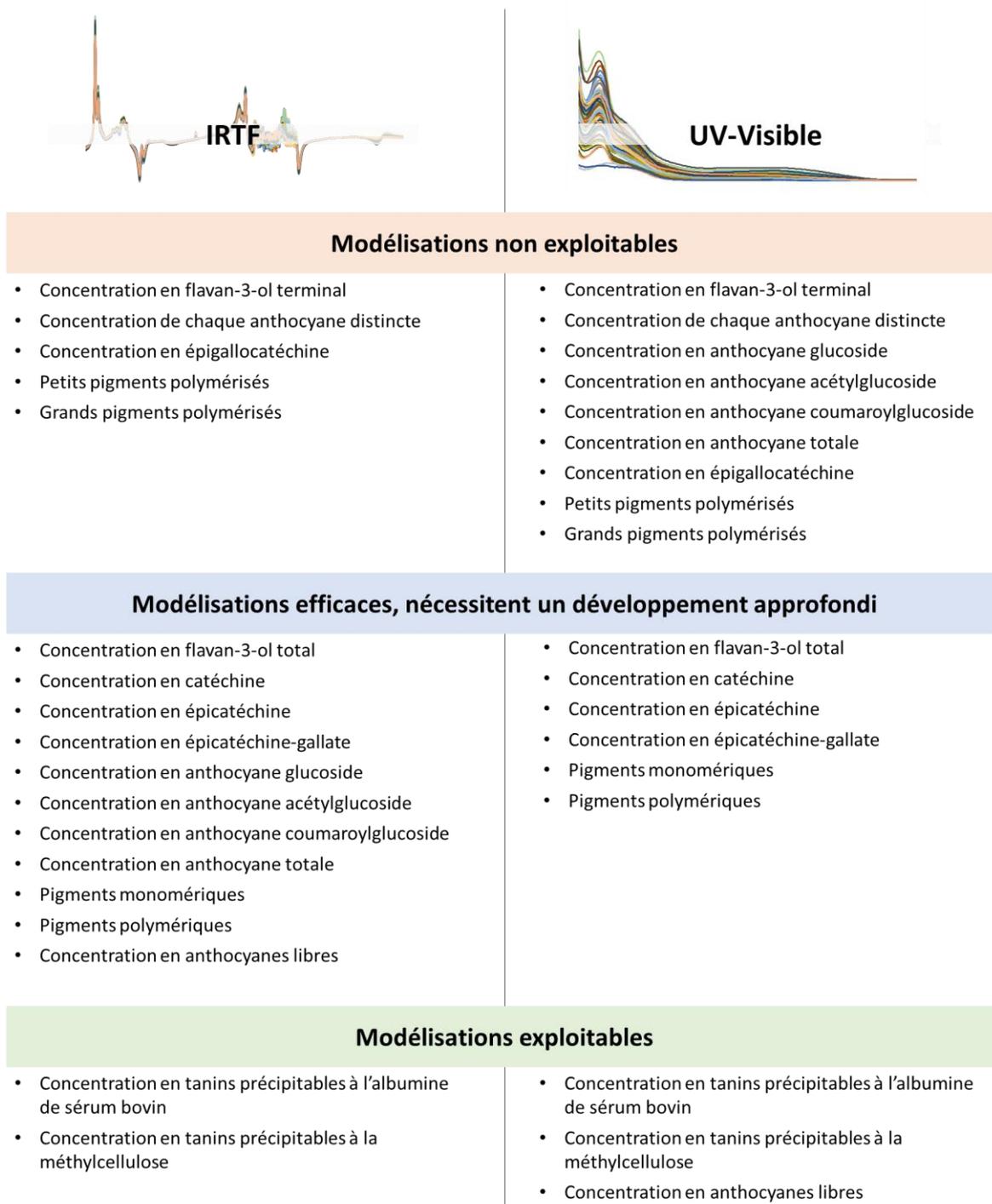


Figure 74 : récapitulatif de la modélisation des anthocyanes et des tanins pendant la fermentation par IRTF et UV-visible

## **C. Suivi de l'évolution des anthocyanes et des tanins pendant le vieillissement en barriques par IRTF et UV-visible couplé à l'analyse PLS**

Nous avons pu suivre plusieurs paramètres relatifs à la concentration en anthocyanes et en tanins grâce à l'IRTF et l'UV-visible pendant la maturité du raisin et la fermentation. Nous avons alors suivi l'évolution des anthocyanes et des tanins pendant le vieillissement en barriques en fonction du temps. Le but est de déterminer s'il est possible de discriminer une évolution de ces paramètres et marqueurs en fûts de chêne comparativement à une évolution en bouteilles. De plus, nous avons voulu déterminer si les modèles construits sur les échantillons de la fermentation étaient adaptables au vieillissement en barriques, où s'il était nécessaire de construire de nouveaux modèles. Pour cela, nous avons suivi quatre vins différents (cépage Merlot) pendant la campagne de vinification de 2018. Le début de l'étude correspond au premier jour de la mise en barriques, après la fermentation malolactique, et un prélèvement d'échantillon a été effectué tous les 2 mois jusqu'à 8 mois d'élevage, et comparé au même vin mis en bouteilles et pour un même temps de prélèvement.

### **1. Comparaison de l'évolution des anthocyanes et des tanins pendant l'élevage en barriques et en bouteilles**

Dans un premier temps, nous avons étudié les paramètres relatifs à l'évolution des anthocyanes et des tanins afin de mettre en évidence des différences caractéristiques entre l'élevage en barriques et en bouteilles.

#### ***a) Évolution des anthocyanes pendant l'élevage en barriques et en bouteilles***

L'évolution de la concentration et de la composition en anthocyanes moléculaires pendant le vieillissement en barriques et en bouteilles est présentée sur la Figure 75.

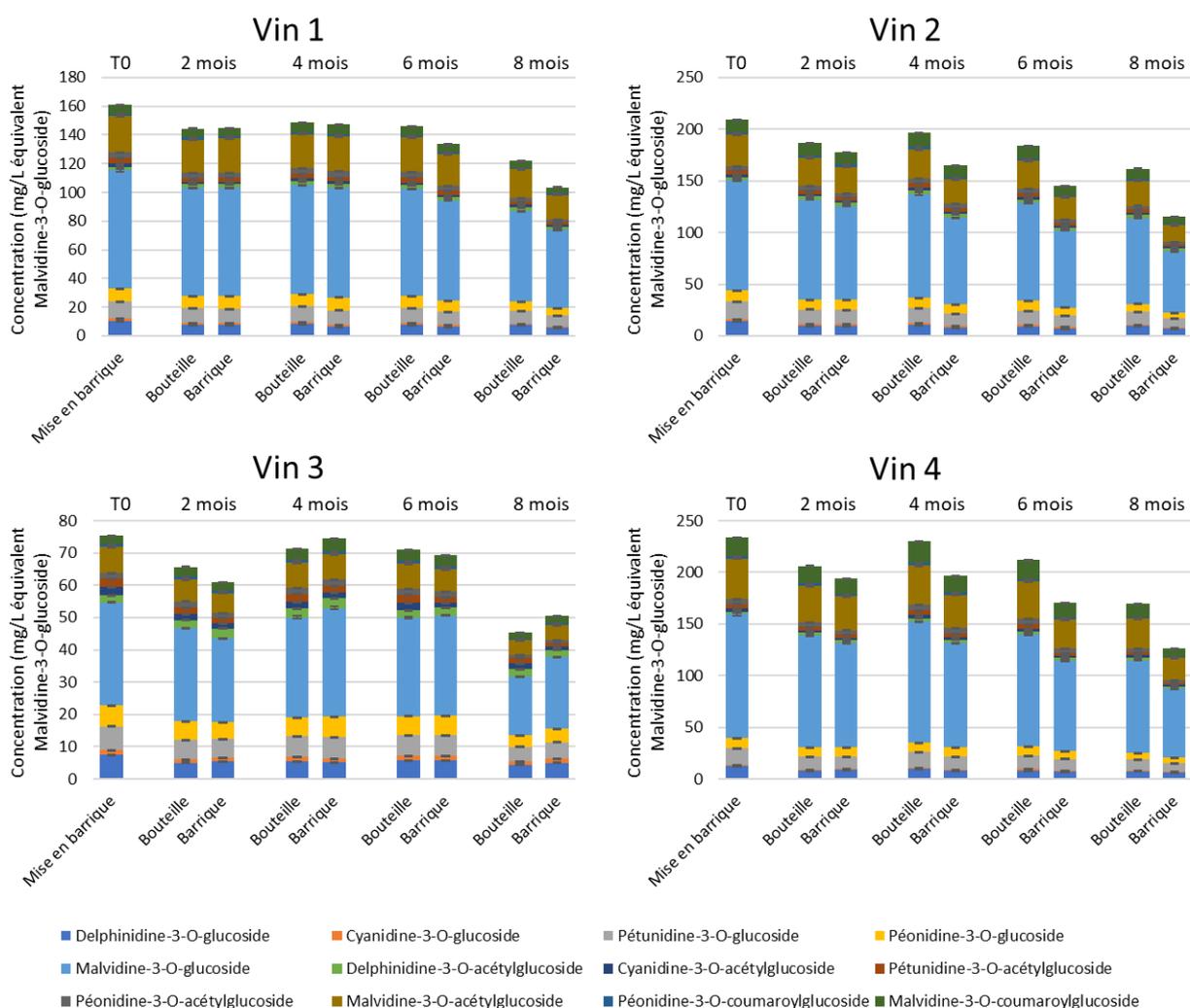


Figure 75 : Evolution de la concentration et de la composition en anthocyanes moléculaires de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles

Nous observons une diminution globale de la concentration en anthocyane au cours du vieillissement. La composition en anthocyanes moléculaires, quant à elle, n'évolue pas de manière caractéristique et les proportions de chaque anthocyane apparaissent identiques tout au long de l'étude. Si l'on compare le vieillissement en barriques et en bouteilles, nous constatons une diminution plus importante de la concentration en anthocyanes pendant l'élevage en barriques. Pour le vin 1, cette différence apparaît distinctement après 6 mois, et est accentuée après 8 mois. Nous observons à cette date un écart de concentration en anthocyanes de 18,50 mg/L équivalent malvidine-3-O-glucoside. Pour le vin 2, ce contraste apparaît directement après 4 mois d'élevage, et à la fin du suivi, l'écart est de 46,70 mg/L équivalent malvidine-3-O-glucoside. Le vin 3 n'apparaît pas sujet à cette variation. Pour le vin 4, cette différence apparaît également après 4 mois, et nous observons un écart entre barriques et bouteilles

après 8 mois de 42,53 mg/L équivalent malvidine-3-*O*-glucoside. Les résultats obtenus indiquent clairement une différence d'évolution entre un vieillissement en barriques et en bouteilles.

Afin de poursuivre l'étude des anthocyanes, nous avons analysé les différences de proportion entre les pigments monomériques et les pigments polymériques présents. Les résultats obtenus sont présentés sur la Figure 76.

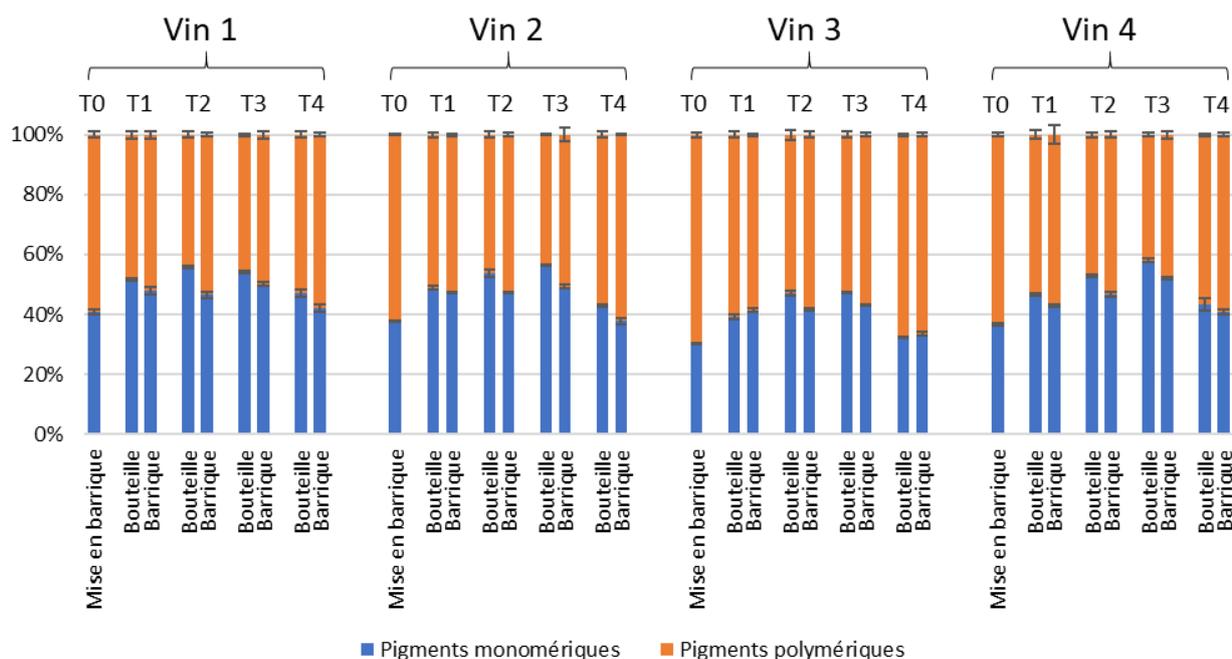


Figure 76 : Evolution de la proportion en pigments monomériques et polymériques de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles (T1 : 2 mois, T2 : 4 mois, T3 : 6 mois, T4 : 8 mois)

Une évolution caractéristique est observée sur tous les vins suivis. Nous constatons une augmentation linéaire de la proportion en pigments monomériques de la mise en barriques jusqu'à un élevage de 6 mois. Entre 6 et 8 mois, cette proportion diminue fortement. Concernant la comparaison entre le vieillissement en barriques et en bouteilles, nous remarquons une proportion en pigments monomériques plus importante pour les vins élevés en bouteilles. Cette différence est observable sur tous les vins suivis à tous les stades de l'élevage, à l'exception du vin 4 après 8 mois, où les proportions sont inversées. Cette disparité peut être attribuée à l'apport d'oxygène plus important pour l'élevage en barriques qui influence de manière significative l'évolution des anthocyanes, en favorisant la polymérisation (355–357).

## *b) Évolution des tanins pendant l'élevage en barriques et en bouteilles*

L'évolution de la concentration en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose pendant le vieillissement en barriques et en bouteilles est présentée sur la Figure 77.

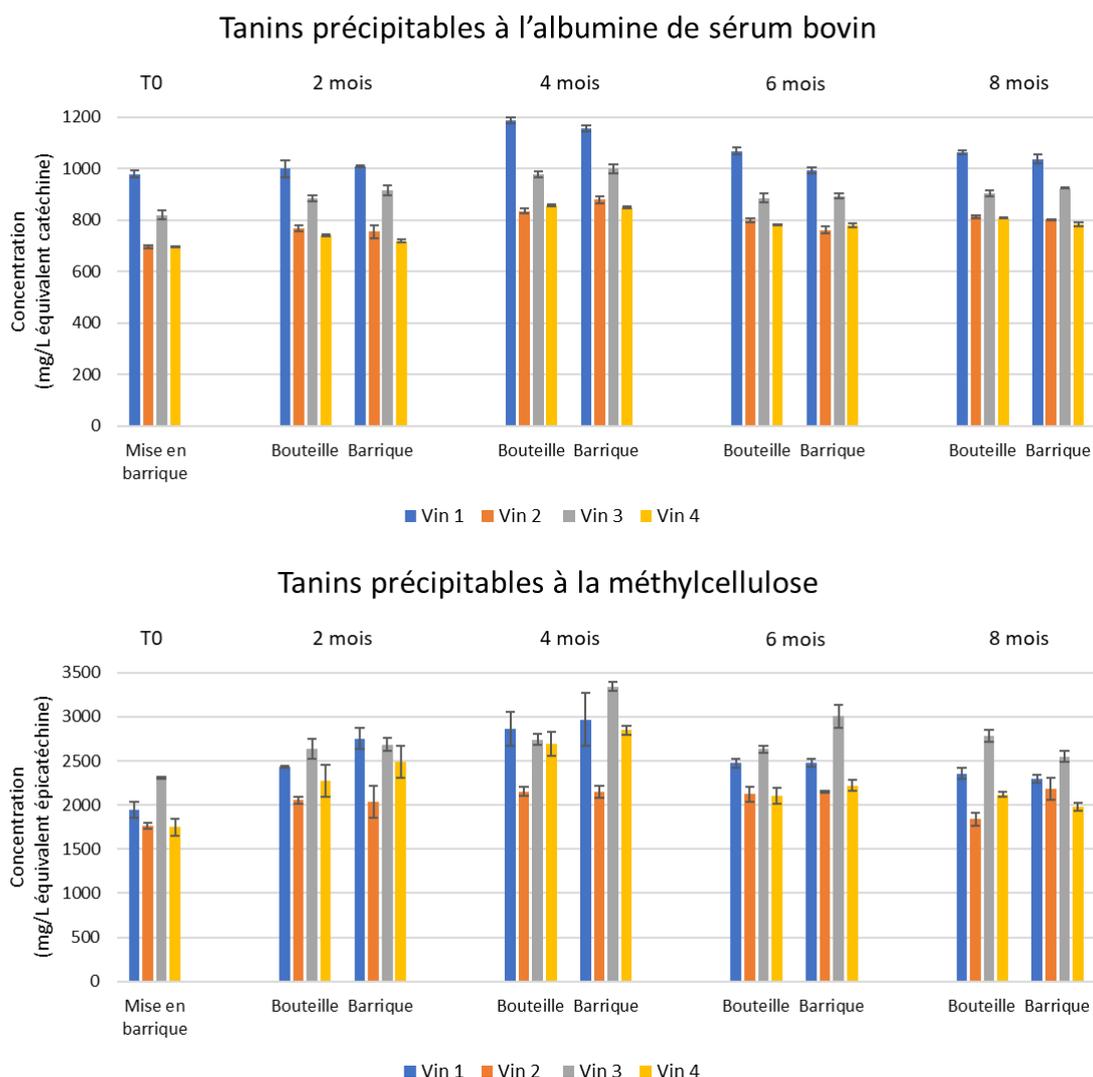


Figure 77 : Evolution de la concentration en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles

Contrairement à l'évolution de la concentration en anthocyanes, les tanins n'évoluent que très peu pendant les 8 premiers mois de l'élevage. Nous constatons cependant une augmentation des tanins précipitables à partir de la mise en barriques et jusqu'à 4 mois, ce qui indique une augmentation de la réactivité des tanins. Cette évolution est facilement observable pour le vin 1 suivi par précipitation

à l'albumine de sérum bovin. Si l'on s'intéresse à l'élevage en barriques, la concentration initiale en tanins et de 979,67 mg/L équivalent catéchine, et augmente jusqu'à 1156,79 mg/L équivalent catéchine après 4 mois.

En comparant le vieillissement en barriques et en bouteilles, nous n'observons pas de différences notables. Pour les tanins précipitables à l'albumine de sérum bovin, les concentrations apparaissent similaires entre l'évolution en barriques et en bouteilles quel que soit le vin étudié. L'on constate le même phénomène mais de façon moins importante pour les tanins précipitables à la méthylcellulose, et les écarts-types apparaissent plus importants ce qui peut expliquer cette différence non significative. Afin d'approfondir cette étude des tanins, l'évolution du degré de polymérisation moyen est présentée sur la Figure 78.

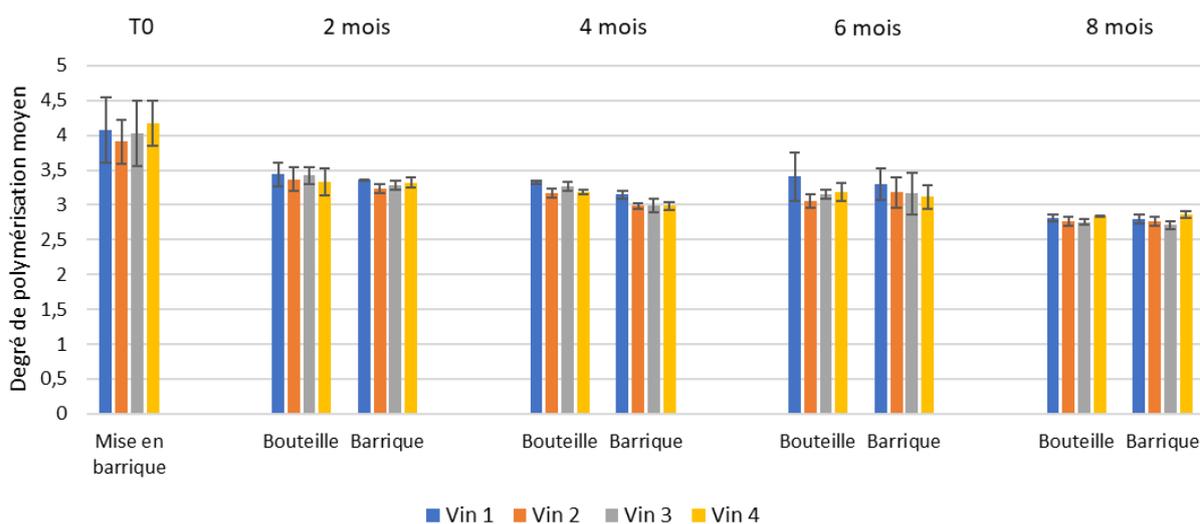


Figure 78 : Evolution du degré de polymérisation moyen de 4 vins de Merlot pendant le vieillissement en barriques et en bouteilles

Une diminution générale du degré de polymérisation moyen est observable pour tous les vins suivis. Les 4 vins montrent des résultats similaires, avec un degré de polymérisation moyen aux alentours de 4 pour la mise en barriques, avec une diminution progressive pour atteindre une valeur inférieure à 3 après 8 mois d'élevage. Cette diminution peut expliquer l'augmentation de la réactivité des tanins observée grâce au suivi des tanins précipitables à l'albumine de sérum bovin. Les chaînes polymériques de flavan-3-ols qui étaient jusque-là trop longues pour précipiter avec une protéine, se retrouvent réduites et précipitent avec l'albumine de sérum bovin. Si nous observons bien une évolution des tanins pendant l'élevage, nous ne distinguons aucune différence entre le vieillissement en barriques et en bouteilles.

## 2. Discrimination du vieillissement en barriques et en bouteilles par IRTF et UV-visible

Nous avons constaté précédemment une évolution différente des anthocyanes en fonction du vieillissement en barriques et en bouteilles. De plus nous avons montré lors de l'étude de la fermentation qu'il était possible de doser les anthocyanes en utilisant l'IRTF ou l'UV-visible couplé à l'analyse PLS. Nous avons donc cherché la possibilité d'utiliser ces analyses spectroscopiques dans la discrimination du vieillissement des vins.

Le modèle PLSDA (analyse discriminante PLS) pour l'UV-visible est présenté sur la Figure 79.

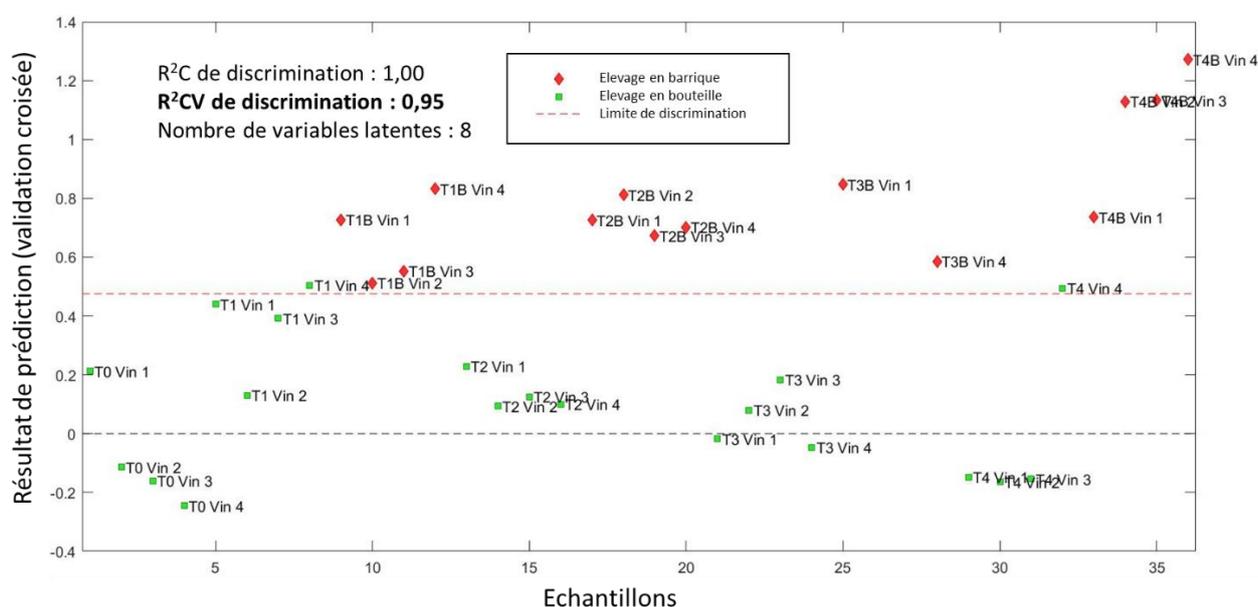


Figure 79 : Analyse PLSDA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des spectres UV-visible (R<sup>2</sup>C : coefficient de régression de calibration, R<sup>2</sup>CV : coefficient de régression de validation croisée, T0 : mise en barriques, T1 : 2 mois, T2 : 4 mois, T3 : 6 mois, T4 : 8 mois)

L'UV-visible montre de très bons résultats pour la discrimination entre le vieillissement en barriques et en bouteilles. Seuls deux échantillons apparaissent mal prédits : le vin 1 en bouteilles après un vieillissement de 2 mois, et le vin 4 en bouteilles après un vieillissement de 8 mois. Nous remarquons que plus le temps d'élevage augmente, plus les échantillons apparaissent s'éloigner de la limite de discrimination. Le modèle créé devient donc de plus en plus sélectif lorsque les vins sont élevés depuis plus longtemps, ce qui implique que la différence d'évolution est bien à l'origine de cette discrimination. Cependant, concernant notre jeu de données, certains échantillons montrent des difficultés à être prédits correctement, malgré une évolution avancée. C'est notamment le cas du vin 4 en bouteilles après un vieillissement de 8 mois qui est mal prédit. Nous constatons qu'il s'agit du

même échantillon qui montrait des proportions en tanins monomériques et polymériques inversées par rapport à la tendance générale (Figure 76). Il est possible que cette bouteille ait été soumise à des concentrations en oxygène plus importante, ce qui pourrait expliquer cette évolution déviante. Cette erreur de prédiction permet cependant de mettre en évidence le rôle prépondérant de l'évolution des anthocyanes, et donc de l'évolution de la couleur, dans la discrimination entre élevage en barriques et en bouteilles.

Les résultats obtenus grâce à l'IRTF sont présentés sur la Figure 80 avec modèle PLS-DA.

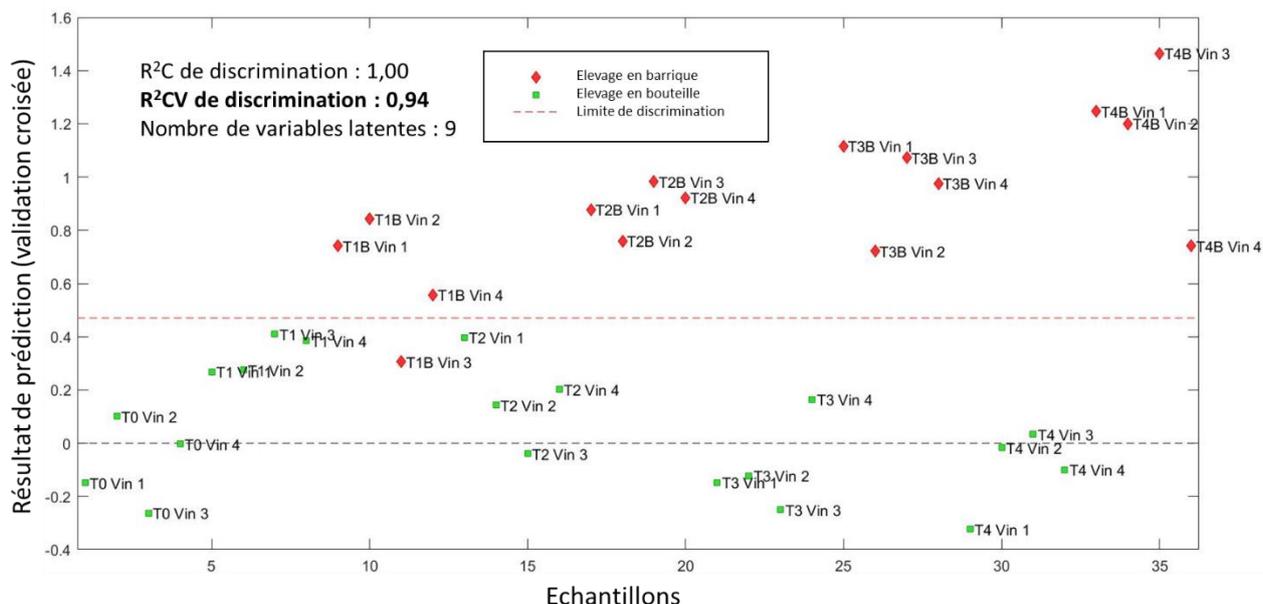


Figure 80 : Analyse PLS-DA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des spectres IRTF ( $R^2C$  : coefficient de régression de calibration,  $R^2CV$  : coefficient de régression de validation croisée, T0 : mise en barriques, T1 : 2 mois, T2 : 4 mois, T3 : 6 mois, T4 : 8 mois)

Avec cette méthode spectroscopique, seul un échantillon est mal prédit : le vin 3 en barriques après un vieillissement de 2 mois. De plus, nous constatons qu'à partir de 4 mois, tous les échantillons sont prédits correctement. À mesure de l'avancée de l'élevage, les résultats de prédiction s'éloignent de la limite de discrimination et indiquent donc une séparation plus prononcée. L'IRTF apparaît donc tout à fait adapté pour discriminer une différence d'élevage pendant le vieillissement.

Considérant les bons résultats obtenus pour la discrimination entre le vieillissement en barriques et en bouteilles avec l'IRTF et l'UV-visible, nous avons recherché la possibilité que les deux méthodes spectroscopiques soient complémentaires et permettent une meilleure séparation. Le modèle PLS-DA pour la combinaison IRTF-UV-visible est présenté en Figure 81.

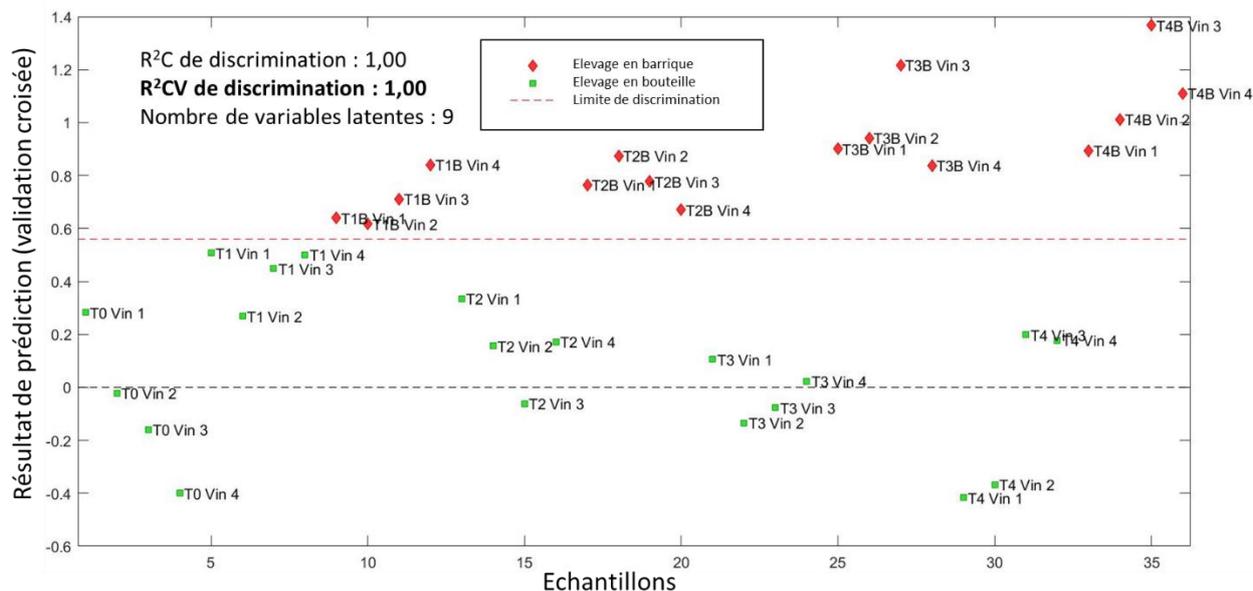


Figure 81 : Analyse PLSDA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des spectres IRTF-UV-visible (R<sup>2</sup>C : coefficient de régression de calibration, R<sup>2</sup>CV : coefficient de régression de validation croisée, T0 : mise en barriques, T1 : 2 mois, T2 : 4 mois, T3 : 6 mois, T4 : 8 mois)

Nous observons avec cette combinaison spectrale une séparation parfaite. Tous les échantillons sont prédits correctement, et la discrimination augmente à mesure de l'élevage. L'IRTF et l'UV-visible apparaissent donc complémentaires, et leur utilisation combinée permet de profiter des avantages des deux méthodes tout en diminuant leurs inconvénients (informations incomplètes, bruits de fond, absorption de composés impactant la mesure).

Considérant le rôle important des concentrations et de l'évolution des anthocyanes dans la discrimination des deux méthodes d'élevage, nous avons recherché s'il était possible de séparer les échantillons grâce aux analyses de références (anthocyanes moléculaires, pigments monomériques et polymériques). Le modèle PLSDA obtenu est présenté sur la Figure 82.

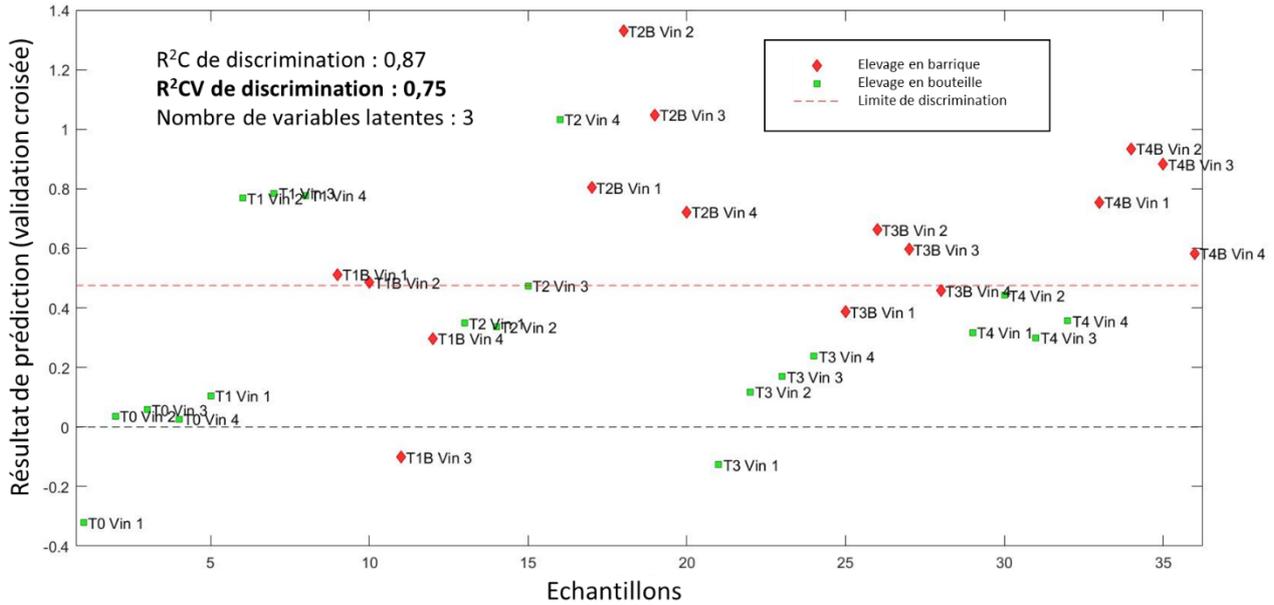


Figure 82 : Analyse PLSDA de l'élevage en barriques et bouteilles de 4 vins de Merlot en fonction des analyses en anthocyanes moléculaires, pigments monomériques et polymériques ( $R^2C$  : coefficient de régression de calibration,  $R^2CV$  : coefficient de régression de validation croisée, T0 : mise en barriques, T1 : 2 mois, T2 : 4 mois, T3 : 6 mois, T4 : 8 mois)

La discrimination apparaît beaucoup moins importante qu'avec les spectres IRTF et UV-visible. Cependant, nous constatons qu'après 8 mois d'élevage, les échantillons sont correctement répartis. Si ces résultats démontrent l'importance des concentrations en anthocyanes et de leur évolution pour distinguer le type d'élevage, ces informations seules ne sont pas suffisantes pour correctement discriminer le vieillissement en barriques et en bouteilles. Une étude similaire a été produite pour utiliser l'IRTF afin de prédire le temps d'élevage des vins rouges et avance l'altération de la couleur comme principal facteur de la modélisation (297). Nous avons pu approfondir les constatations de ce travail en mettant en évidence l'impact important de l'évolution et de la concentration en anthocyanes pour la discrimination de l'élevage. Néanmoins, ces informations seules ne permettent pas une discrimination efficace, et semblent indiquer que d'autres composés influencent la séparation. Il serait désormais intéressant de poursuivre cette étude en intégrant le dosage de composés spécifiques au bois tels que les ellagitanins.

### 3. Prédiction des concentrations en anthocyanes et tanins de vins de Merlot pendant l'élevage à l'aide des modèles construits pour la fermentation

Nous avons voulu évaluer si les modèles de prédiction de la fermentation étaient adaptables à l'élevage des vins, où s'il était nécessaire de construire de nouveaux modèles. Pour cela, nous avons sélectionné les meilleurs modèles obtenus pour l'IRTF (somme des anthocyanes moléculaires, tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose) et l'UV-visible (anthocyanes libres, tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose), et nous avons cherché à prédire ces paramètres pendant l'élevage. Les résultats obtenus sont présentés dans le Tableau 31. Nous avons décidé de calculer deux écarts prédictifs résiduels (RPD P) différents. Le premier se base sur l'écart-type du suivi de la fermentation (RPD P fermentation), et permet d'évaluer si les échantillons prédits s'intègrent dans la variabilité de la calibration. Le second se base sur l'écart-type de l'élevage (RPD P élevage), et permet d'évaluer si la prédiction est assez robuste pour observer la variabilité du paramètre pendant l'élevage.

Paramètre	R <sup>2</sup> P	RMSEP	RPD P fermentation	RPD P élevage	VL
<b>IRTF</b>					
Total anthocyanes (mg/L)	0,01	254,14	0,42	0,21	18
BSA (mg/L)	0,37	268,43	0,86	0,54	10
MCP (mg/L)	0,26	356,71	1,82	1,16	10
<b>UV-visible</b>					
Anthocyanes libres (mg/L)	0,62	55,54	2,43	1,61	10
BSA (mg/L)	0,46	333,25	0,69	0,43	15
MCP (mg/L)	0,32	381,22	1,70	1,08	11

Tableau 31 : Prédiction des anthocyanes et des tanins durant l'élevage pour les vins de Merlot étudiés par les modèles IRTF et UV-visible construits avec les échantillons de moûts et vins du suivi de la fermentation (BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Pour l'IRTF, nous constatons que la prédiction de la somme des anthocyanes moléculaires (RPD P fermentation : 0,42) et des tanins précipitables à l'albumine de sérum bovin (RPD P fermentation : 0,86) n'est pas assez robuste pour prédire les concentrations pendant l'élevage. Pour les tanins précipitables à la méthylcellulose, nous obtenons un RPD P fermentation de 1,82, les échantillons peuvent donc être prédits avec une erreur acceptable. Cependant, le RPD P élevage de 1,16 n'est pas assez élevé pour permettre d'évaluer efficacement la variabilité étudiée pendant le vieillissement, il

apparaît donc nécessaire de développer de nouveaux modèles afin d'obtenir des valeurs fiables en utilisant l'IRTF.

Concernant l'UV-visible, nous observons des résultats semblables pour les tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose. Pour les tanins précipitables à l'albumine de sérum bovin, nous avons un RPD P fermentation de 0,69, très insuffisant pour envisager une prédiction correcte. Pour les tanins précipitables à la méthylcellulose, le RPD P fermentation obtenu est de 1,70, ce modèle pourrait donc être envisageable, mais le RPD P élevage n'est que de 1,08 et n'apparaît donc pas adapté à la variabilité plus faible observée pendant l'élevage. Pour la prédiction des anthocyanes libres, le RPD P fermentation de 2,43 montre une très bonne robustesse. De plus, le RPD P élevage de 1,70 est suffisamment élevé pour être considéré comme fiable. Le modèle de prédiction des anthocyanes libres est présenté sur la Figure 83.

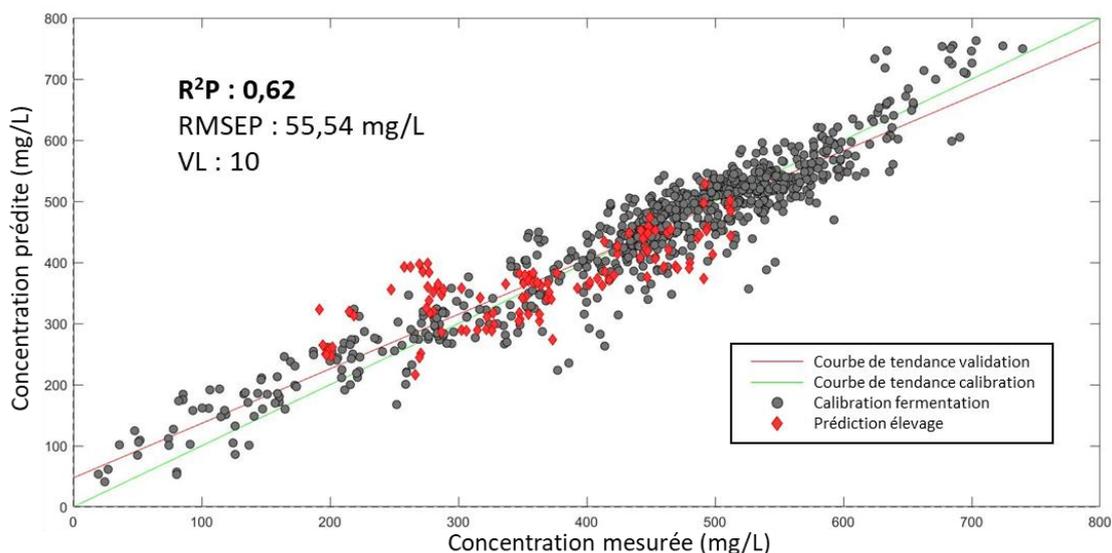


Figure 83 : Prédiction des anthocyanes libres durant l'élevage pour les vins de Merlot étudiés par le modèle UV-visible construit avec les échantillons de moûts et vins du suivi de la fermentation (R<sup>2</sup>P : Coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, VL : nombre de variables latentes)

Les échantillons prédits de l'élevage apparaissent bien intégrés à la calibration obtenue avec les échantillons du suivi de la fermentation. Nous observons peu de déviations des échantillons étudiés, ce modèle apparaît donc tout à fait adapté pour suivre l'évolution des anthocyanes libres pendant l'élevage.

### ***a) Conclusions et perspectives de la prédiction des anthocyanes et des tanins par IRTF et UV-visible pendant l'élevage***

Nous avons étudié la différence entre un élevage en barriques et en bouteilles, afin de rechercher la possibilité de les discriminer à l'aide de l'IRTF et l'UV-visible. Nous avons obtenu une très bonne séparation entre les échantillons des deux élevages à l'aide des spectres IRTF et UV visible. En utilisant la combinaison des deux méthodes spectroscopiques, nous avons montré qu'il était possible d'améliorer cette discrimination, et nous avons obtenu un modèle capable de séparer parfaitement les échantillons suivis en fonction du type de vieillissement. Considérant l'évolution différente de la concentration des anthocyanes et ses dérivés en fonction de l'élevage, nous avons recherché si ces paramètres seul permettaient de différencier les différents types d'échantillons. Nous avons mis en évidence que si ces concentrations étaient des facteurs importants dans la discrimination, elles ne suffisaient pas pour obtenir une prédiction aussi efficace que celles obtenues avec l'analyse spectrale. D'autres paramètres, tels que les composés du bois comme les ellagitanins ou encore l'oxygène dissous doivent influencer la mesure spectrale et permettre une meilleure séparation des deux types d'élevage étudiés. Afin d'approfondir ce travail, il serait intéressant d'ajouter des échantillons et de mesurer ces paramètres pour identifier les composés responsables de cette discrimination.

Nous avons cherché s'il était possible de prédire les concentrations en anthocyanes et en tanins pendant l'élevage avec les modèles construits avec les échantillons du suivi de la fermentation. Nous avons mis en évidence des résultats intéressants avec l'IRTF et l'UV-visible pour la prédiction des tanins précipitables à la méthylcellulose. Cependant, la fiabilité du modèle n'était pas suffisante pour prédire correctement cette concentration pendant l'élevage, à cause de la faible variabilité observée. Nous avons démontré que le modèle des anthocyanes libres construit à l'aide des spectres UV visibles était tout à fait adaptable pour la prédiction d'échantillons pendant l'élevage. Pour les autres paramètres, il apparaît nécessaire de développer de nouveaux modèles.

### **D. Construction de modèles de prédiction PLS pour le suivi des anthocyanes et des tanins dans les vins fins commerciaux**

Les concentrations en anthocyanes et tanins ont été analysés sur 92 vins fins commerciaux. Nous avons recherché en priorité la diversité, et les vins étudiés sont répartis sur 12 pays différents et 16 millésimes allant de 1997 à 2017. Comme pour le suivi de la maturité et de la fermentation, la

première étape de cette étude consiste à étudier le jeu de données obtenu afin d'évaluer la variabilité des échantillons.

## **1. Analyse du jeu de données**

### ***a) Evaluation de la variabilité des analyses de références***

Sur tous les échantillons, les analyses suivantes ont été effectuées :

- L'analyse de la composition en flavan-3-ols par CLHP-UV-visible après réaction de phloroglucinolyse.
- Le dosage des anthocyanes moléculaires par CLHP-UV-visible.
- L'analyse de l'évolution des anthocyanes, avec le dosage d'Adams-Harbertson
- Le dosage des anthocyanes libres par décoloration au bisulfite de sodium.
- Le dosage des tanins par précipitation à l'albumine de sérum bovin et à la méthylcellulose
- Le dosage des tanins totaux par hydrolyse acide (dosage de Bate-Smith)

Les variabilités de ces analyses sont présentées dans le Tableau 32. Aucun échantillon n'a été écarté.

Paramètre	Moyenne	Min	Max	Ecart-type	CV	Ecart-type Ref
<b>Composition en flavan-3-ols</b>						
Flavan-3-ol (mmol/L)	13,76	0,16	26,60	5,31	39%	0,79
Flavan-3-ol terminal (mmol/L)	5,41	0,00	12,04	2,35	43%	0,62
Catéchine (mmol/L)	3,00	0,00	6,98	1,06	35%	0,24
Epicatéchine (mmol/L)	5,99	0,16	14,02	2,91	49%	0,41
Epigallocatechine (mmol/L)	4,47	0,00	11,95	2,43	55%	0,66
Epicatéchine-gallate (mmol/L)	0,30	0,00	1,23	0,19	62%	0,02
<b>Anthocyanes moléculaires</b>						
Dp-3-O-glc (mg/L)	3,99	0,00	24,74	4,93	124%	0,19
Cy-3-O-glc (mg/L)	0,67	0,00	4,86	0,79	118%	0,11
Pt-3-O-glc (mg/L)	4,30	0,00	31,20	5,30	123%	0,18
Pn-3-O-glc (mg/L)	2,74	0,00	17,48	3,45	126%	0,12
Mlv-3-O-glc (mg/L)	26,49	0,00	116,50	29,95	113%	0,37
Dp-3-O-acglc (mg/L)	1,97	0,00	8,75	2,19	111%	0,10
Cy-3-O-acglc (mg/L)	0,62	0,00	2,41	0,58	93%	0,08
Pt-3-O-acglc (mg/L)	0,97	0,00	6,13	1,31	135%	0,10
Pn-3-O-acglc (mg/L)	0,69	0,00	4,67	0,98	143%	0,08
Mlv-3-O-acglc (mg/L)	6,06	0,00	44,88	9,04	149%	0,15
Pn-3-O-cmglc (mg/L)	0,53	0,00	2,69	0,52	98%	0,06
Mlv-3-O-cmglc (mg/L)	2,32	0,00	11,54	2,45	105%	0,11
Total glucoside (mg/L)	38,20	0,12	173,35	42,08	110%	0,65
Total acétyl (mg/L)	8,33	0,00	52,66	11,52	138%	0,24
Total coumaroyl (mg/L)	2,85	0,00	12,76	2,83	99%	0,14
Total anthocyanes (mg/L)	49,38	0,12	208,20	55,23	112%	0,81
<b>Evolution des anthocyanes</b>						
MP (Absorbance)	1,20	0,03	2,75	0,52	43%	0,06
SPP (Absorbance)	1,35	0,14	3,21	0,58	43%	0,05
LPP (Absorbance)	1,16	0,00	4,84	0,88	76%	0,07
PP (Absorbance)	2,51	0,11	7,37	1,26	50%	0,06
Anthocyanes libres (mg/L)	152,09	12,19	557,32	115,24	76%	4,84
<b>Tanins</b>						
BSA (mg/L)	425,91	0,00	933,81	206,89	49%	15,91
MCP (mg/L)	1594,80	0,00	3842,69	644,26	40%	129,66
Tanins totaux (mg/L)	4010,43	219,72	7748,75	1437,53	36%	297,70

Tableau 32 : Paramètres des analyses de références obtenus sur les vins finis commerciaux (Dp : delphinidine, Cy : cyanidine, Pt :pétunidine, Pn :péonidine, : Mlv : malvidine, glc : glucoside, : acglc :acétylglucoside, cmglc : coumaroylglucoside, MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, CV coefficient de variation)

En observant ces données, nous constatons des concentrations en anthocyanes moléculaires très faibles. Hormis la malvidine-3-O-glucoside, qui compose environ 50 % de la concentration en anthocyanes totales, toutes les autres anthocyanes ont une moyenne de concentration inférieure à

10 mg/L. Il a été démontré précédemment avec le suivi de la maturité et de la fermentation que pour ces composés nous obtenions des erreurs d'interprétation lors de la prédiction, dues à une corrélation avec la somme des anthocyanes moléculaires. En considérant la difficulté de modéliser les faibles concentrations en utilisant l'IRTF et l'UV-visible, il a été décidé de ne pas considérer chaque anthocyane moléculaire indépendante, mais plutôt la somme des anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales. Certaines des variétés étudiées, telles que le Pinot noir, montrent une déficience en anthocyanes acétylées, il apparaît donc important de pouvoir suivre ces paramètres. Pour la somme des anthocyanes moléculaires, nous obtenons la plus forte variabilité observée, avec un coefficient de variation de 112 %. Etant donné la différence des millésimes étudiés et la stabilité relative des anthocyanes dans le temps, cette variabilité peut être attribuée à l'évolution des anthocyanes pendant le vieillissement. Pour valider cette hypothèse, nous avons présenté la concentration de la somme des anthocyanes en fonction du millésime sur la Figure 84

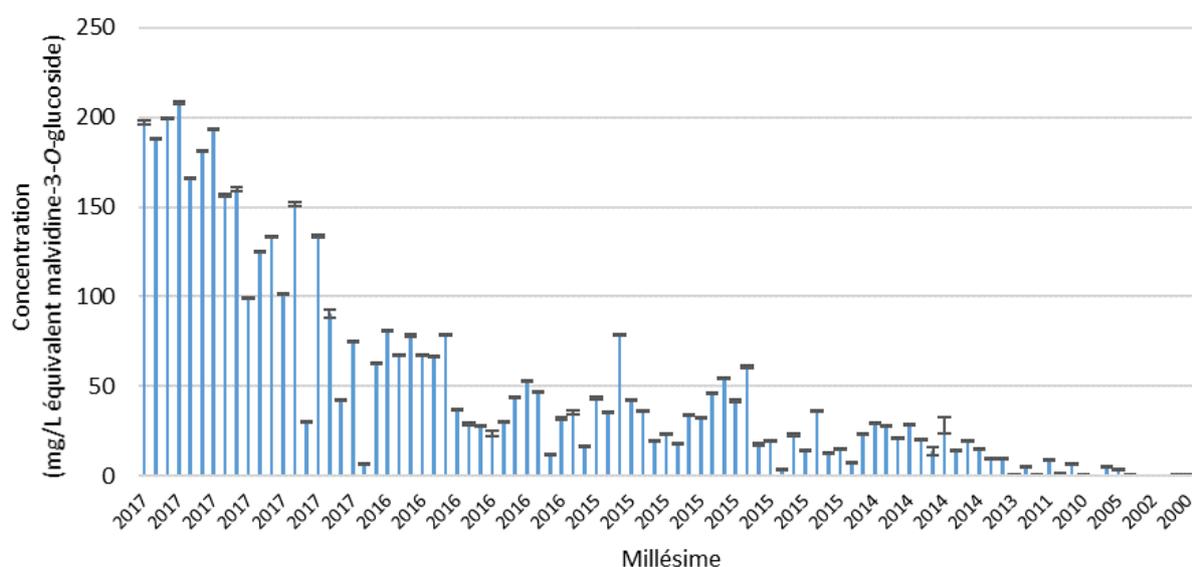


Figure 84 : Concentration de la somme des anthocyanes moléculaires en fonction du millésime sur les vins fins commerciaux

Nous constatons clairement une diminution de la concentration en anthocyanes moléculaires pendant le vieillissement. Cependant, nous observons aussi d'importantes différences entre vins d'un même millésime, et il n'est pas possible d'attribuer une valeur de concentration pour un millésime donné. Pouvoir modéliser la concentration en anthocyanes grâce à l'IRTF et l'UV-visible apparaît donc idéale pour estimer la composition phénolique d'un vin, quel que soit le millésime étudié.

Nous observons également une bonne variabilité pour l'évolution des anthocyanes. Les pigments monomériques et les petits pigments polymérisés présentent le même coefficient de variation (43 %).

Les grands pigments polymérisés, quant à eux, montrent une variabilité accrue avec un coefficient de variation de 76 %. Leurs proportions en fonction du millésime sont présentées sur la Figure 85.

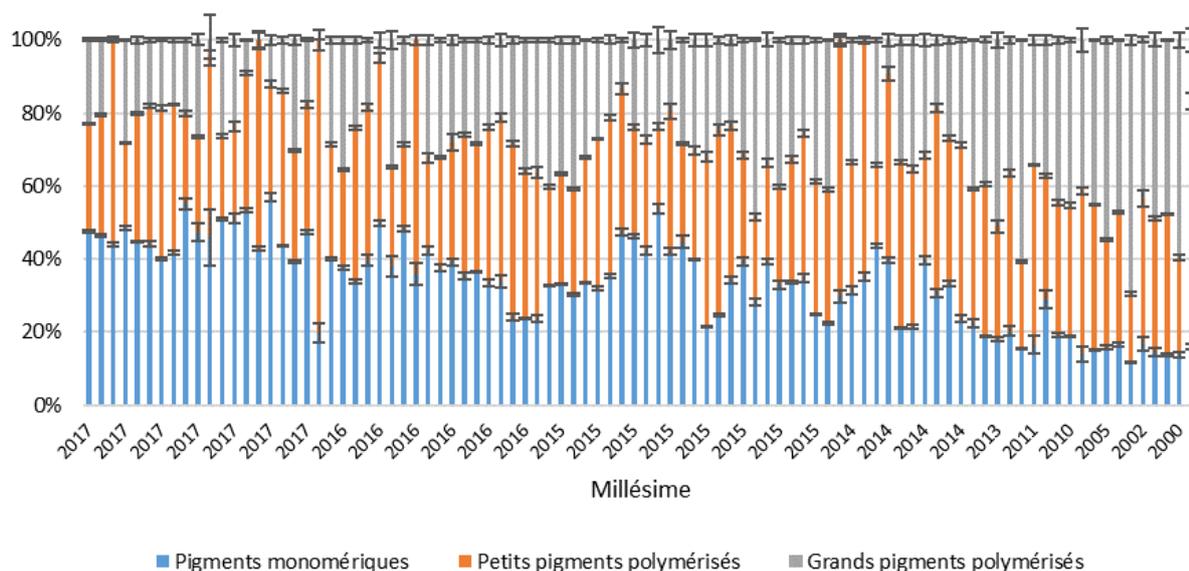


Figure 85 : Proportion en pigments monomériques, petits pigments polymérisés et grands pigments polymérisés en fonction du millésime sur les vins finis commerciaux

Le même constat que pour les anthocyanes moléculaires peut être fait. Nous observons une diminution des pigments monomériques pendant le vieillissement du vin, mais nous constatons néanmoins que nous ne pouvons pas attribuer à un millésime des proportions en pigments définies. Le pourcentage des grands pigments polymérisés augmente à mesure que le pourcentage de pigments monomériques diminue, et les petits pigments polymérisés ne semblent pas impactés par l'évolution du vieillissement.

Si l'on s'intéresse aux coefficients de variation pour les paramètres des tanins condensés, qui mesurent des informations proches, nous observons des différences significatives. Les tanins totaux mesurés par hydrolyse acide indiquent le coefficient de variation le plus faible, avec 36 %. Cette valeur peut s'expliquer par le fait que cette méthode mesure la concentration en tanins indépendamment de leur réactivité. A contrario, les tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose indiquent une plus grande variabilité, avec un coefficient de variation respectif de 49 % et 40 %. Si le dosage à la méthylcellulose se rapproche de la variabilité des tanins totaux par hydrolyse acide, nous constatons que pour le dosage à l'albumine de sérum bovin, celle-ci est plus importante. Cette méthode basée sur une protéine mesure de manière plus spécifique la réactivité des tanins, et il apparaît donc que celle-ci varie plus fortement sur le jeu d'échantillons que la concentration totale en tanins. Plusieurs études ont démontré que la réactivité des tanins est impactée par l'évolution du degré

de polymérisation moyen (3,246). Nous avons donc cherché si le degré de polymérisation moyen suivait une évolution caractéristique en fonction du millésime (Figure 86).

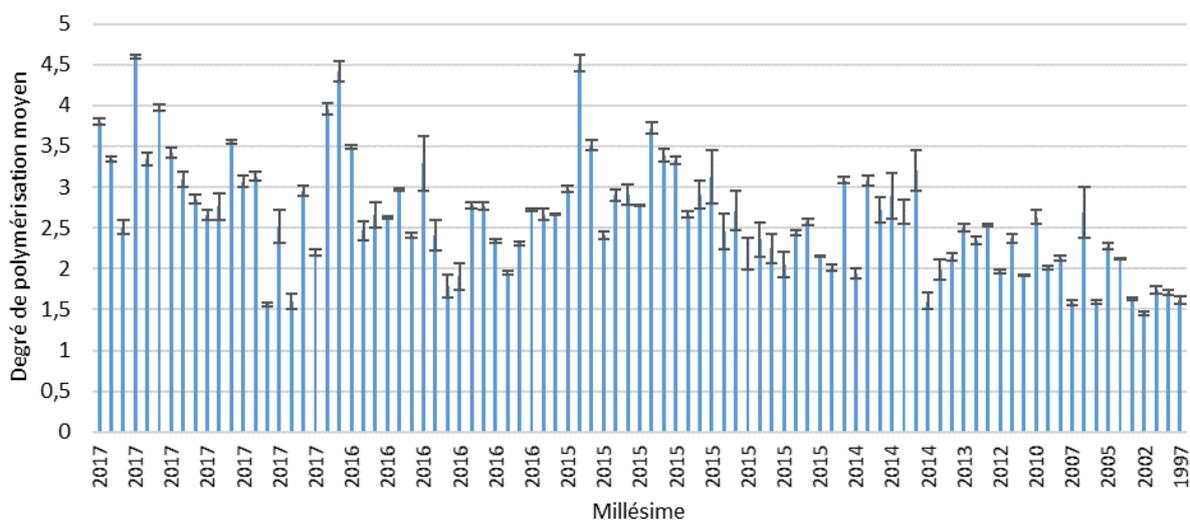


Figure 86 : Evolution du degré de polymérisation moyen en fonction du millésime sur les vins finis commerciaux

Nous constatons une tendance générale à la diminution du degré de polymérisation moyen avec l'avancée du vieillissement. Cependant, contrairement à l'évolution des anthocyanes, elle apparaît beaucoup moins importante. Le degré de polymérisation moyen est impacté par des réactions de condensation et de dépolymérisation, mais aussi par le cépage étudié et les proportions en tanins de pépins et de pellicules (118,119,179). Ainsi, la grande diversité du jeu d'échantillons peut expliquer la variabilité observée.

De manière générale, nous observons de très fortes variabilités pour tous les paramètres étudiés, avec un coefficient de variation minimum de 35 % pour la concentration en unités catéchines et un coefficient de variation maximum de 112 % pour la somme des anthocyanes moléculaires. Le jeu d'échantillons étudié apparaît adapté pour la modélisation des anthocyanes et des tanins. Afin de compléter cette étude, nous avons réalisé des analyses en composantes principales (ACP) pour déterminer les paramètres corrélés entre eux, ainsi que la dispersion des échantillons en fonction des analyses de références et des analyses spectrales.

## *b) Analyse en composantes principales (ACP) des analyses de références et des données spectrales*

Les analyses en composantes principales (ACP) des individus des analyses de références, des variables des analyses de références, des individus des spectres UV-visible et IRTF sont présentées sur la Figure 87. Le millésime a été utilisé comme variable externe.

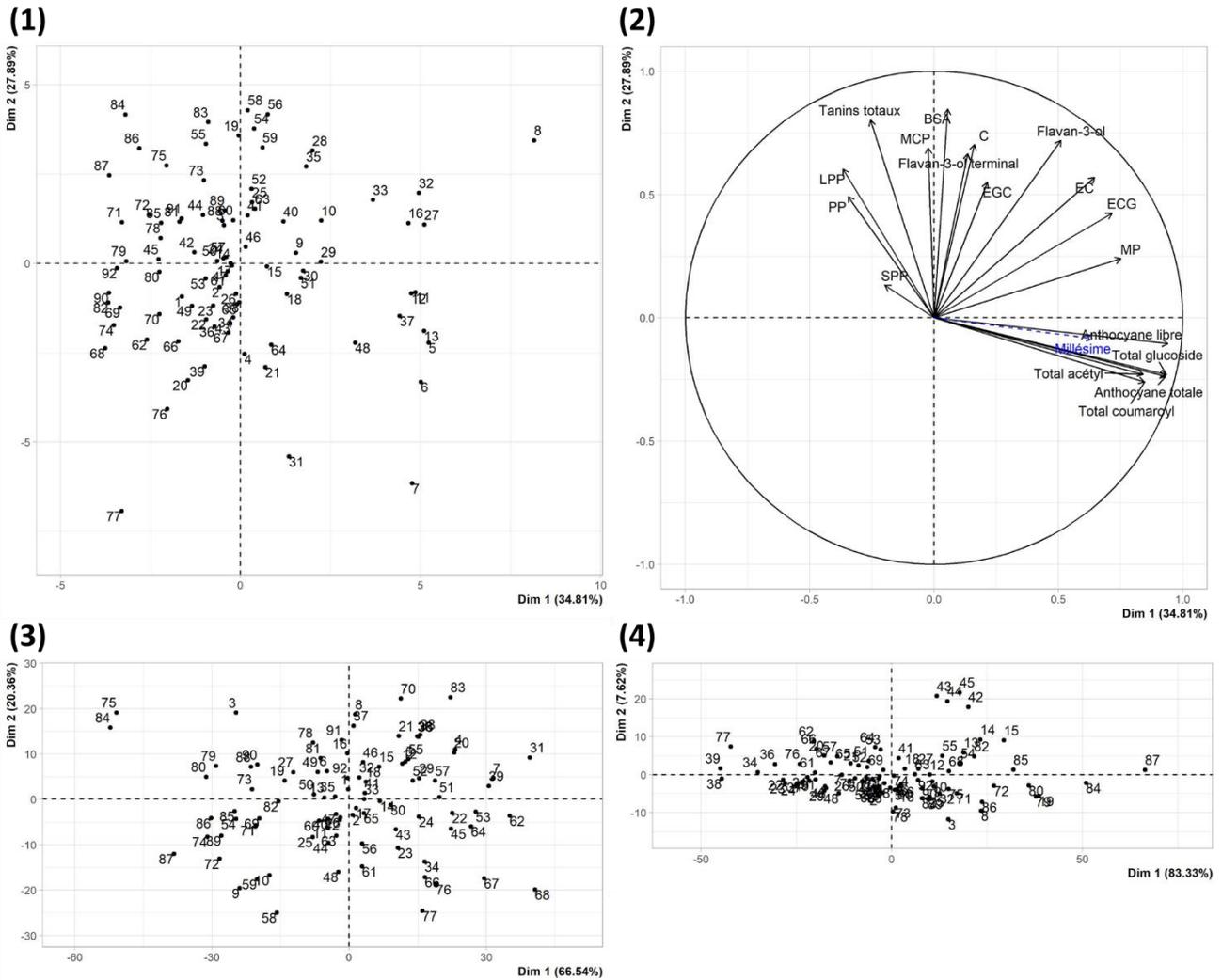


Figure 87 : Analyses en composantes principales, dimension 1 et 2 : (1) individus des analyses de références, (2) variables des analyses de références, (3) individus des spectres IRTF, (4) individus des spectres UV-visible (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, C : catéchine, EC : épicatechine, EGC : épigallocatechine, ECG : épicatechine-gallate)

Pour les 3 ACP, nous nous sommes intéressés aux deux premières dimensions, qui expliquent la plus grande proportion de variation du jeu de données. Pour l'ACP des analyses de références, ces deux

dimensions expliquent 62,70 % de la variation totale pour l'ACP des analyses de références, 86,90 % de la variation totale pour l'ACP des spectres IRTF, et 90,95 % de la variation totale pour l'ACP des spectres UV-visible. Sur l'ACP des individus pour l'analyse des analyses de références, nous constatons une distribution des vins relativement homogène, avec néanmoins plusieurs échantillons qui se détachent du nuage de points. C'est notamment le cas des échantillons 7, 8, 31, 77, 84 et 87. Si l'on compare avec les ACP des spectres, nous constatons que pour l'UV-visible, 3 de ces échantillons sont en commun, les vins 77, 84 et 87. Pour les spectres IRTF, ces 3 vins apparaissent aussi se détacher du nuage de points, ainsi que l'échantillon 31 et 75. Nous avons cherché à expliquer cette déviance par leur composition phénolique. Il apparaît que les vins 7, 31 et 77 se distinguent par une faible concentration en tanins condensés. A contrario, les vins 75, 84 et 87 indiquent une concentration en tanins plus élevée que la moyenne. Les vins 7 et 8 se démarquent quant à eux par une forte concentration en anthocyanes moléculaires. Ainsi, la répartition étudiée apparaît globalement homogène, avec plusieurs échantillons déviants qui peuvent s'expliquer par leur composition phénolique. Le jeu de données est donc adapté à la conception de modèle PLS et aucun échantillon ne devra être écarté.

Si l'on s'intéresse au graphique des variables des analyses de références, nous constatons que 2 groupes se forment et apparaissent fortement corrélés aux axes des dimensions 1 et 2. Le premier groupe, composé de la somme des anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales, des anthocyanes libres et des pigments monomériques, apparaît fortement corrélé à la dimension 1 et explique la variabilité de la concentration en anthocyanes. Le second groupe, composé des tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose, des tanins totaux, et des concentrations en unités flavan-3-ols terminales, catéchine et épigallocatechine est corrélé avec la dimension 2 et explique la variabilité de la concentration en anthocyanes. Pour les grands pigments polymérisés et la somme des pigments polymérisés, nous constatons une corrélation négative avec les paramètres décrivant les anthocyanes monomères, et indique bien une condensation de ces composés avec les tanins comme cela avait pu être observé sur la Figure 85. De plus, nous observons une corrélation entre le millésime et les anthocyanes monomères qui corrobore la diminution de concentration de celles-ci observée pendant le vieillissement.

Afin d'observer de possibles regroupements dus aux millésimes ou aux cépages en fonction de la composition phénolique, des ACP des analyses de références mettant en évidence ces sous-groupes sont présentés sur la Figure 88.

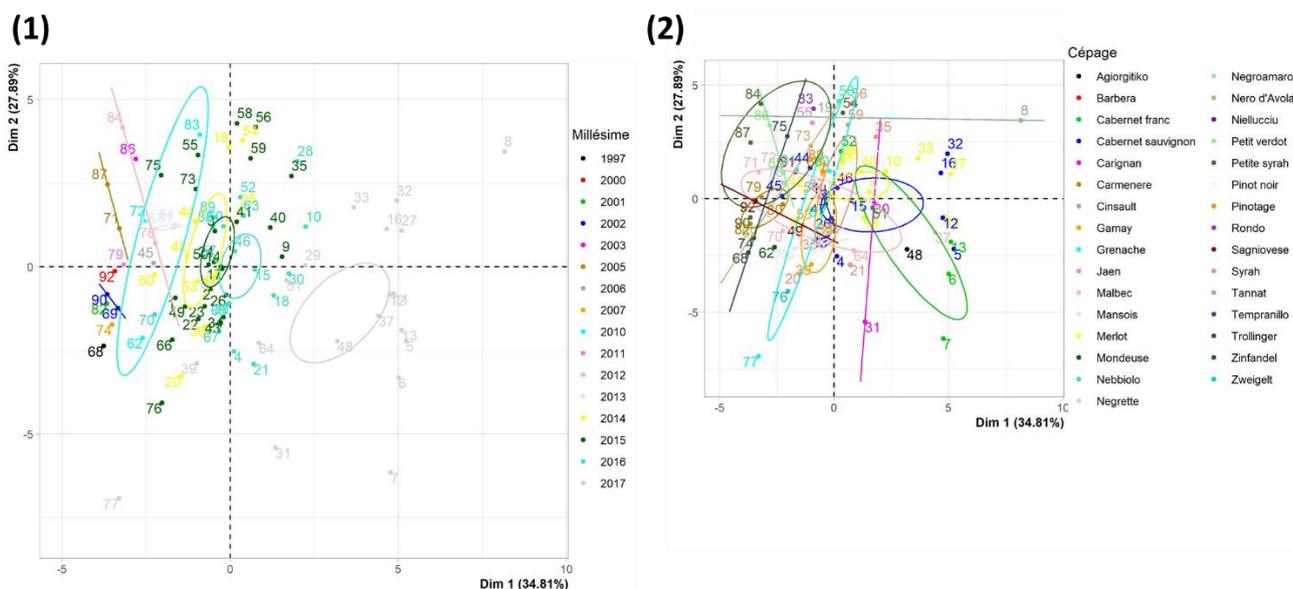


Figure 88 : Analyses en composantes principales des individus des analyses de références en fonction des millésimes (1) et des cépages (2), dimension 1 et 2

En regardant le premier graphique des individus classés par millésime, nous constatons que plusieurs sous-groupes peuvent être construits. En effet, le millésime 2017, plus jeune millésime étudié, se détache clairement des autres échantillons. Ensuite, il est possible de regrouper les échantillons des millésimes 2016, 2015 et 2014. Les vins plus anciens n'apparaissent pas assez différents pour faire un nouveau regroupement. Cette séparation est faite le long de l'axe de la dimension 1, il est alors possible de considérer les différents paramètres contribuant à l'analyse des anthocyanes comme un marqueur du vieillissement du vin.

Si l'on s'intéresse au graphique des individus classés par cépage, nous ne constatons aucun regroupement, avec une superposition des ellipses de confiance. Ainsi, si la composition des polyphénols peut varier d'une variété à l'autre, cette variabilité n'apparaît pas assez importante pour différencier les cépages. Cependant, cette constatation est seulement applicable à ce jeu de données, qui inclut un nombre important de cépages au regard du nombre d'échantillons.

Nous avons pu constater que notre jeu de données montrait une grande diversité, ainsi que très peu d'échantillons déviants qui ne pouvaient être expliqués par leur composition en polyphénols. Il apparaît donc tout à fait adapté à la construction de modèle PLS.

## **2. Construction de modèles de prédiction PLS pour le suivi des anthocyanes et des tanins sur les vins finis commerciaux**

### ***a) Comparaison entre les modèles PLS construits avec les spectres IRTF et UV-visible pour la prédiction des anthocyanes et des tanins sur les vins finis commerciaux***

Grâce aux données des analyses de références recueillies sur les échantillons de vins finis commerciaux, nous avons pu construire des modèles de prédiction en se basant sur les spectres IRTF et UV-visible. Pour valider la modélisation, nous avons utilisé la validation croisée leave-p-out cross-validation, avec 10 sous-groupes. Nous avons également utilisé l'algorithme de Kennard-Stone pour diviser le jeu de données en un groupe de calibration (75 % des échantillons) et de validation externe (25 % des échantillons). Les résultats de modélisation obtenus pour l'IRTF sont présentés dans le Tableau 33, et les résultats obtenus pour l'UV-visible sont présentés dans le Tableau 34.

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Composition en flavan-3-ols</b>								
Flavan-3-ol (mmol/L)	0,80	0,73	0,65	2,88	2,63	1,84	2,02	15
Flavan-3-ol terminal (mmol/L)	0,54	0,42	0,33	1,94	1,78	1,21	1,32	14
Catéchine (mmol/L)	0,69	0,62	0,29	0,69	0,83	1,53	1,27	11
Epicatéchine (mmol/L)	0,84	0,81	0,72	1,28	1,36	2,27	2,14	13
Epigallocatechine (mmol/L)	0,52	0,36	0,36	2,17	1,67	1,12	1,46	15
Epicatéchine-gallate (mmol/L)	0,77	0,66	0,25	0,11	0,15	1,69	1,24	18
<b>Anthocyanes moléculaires</b>								
Total glucoside (mg/L)	0,81	0,74	0,64	19,94	25,26	2,11	1,67	16
Total acétyl (mg/L)	0,79	0,72	0,74	5,41	7,15	2,13	1,61	15
Total coumaroyl (mg/L)	0,77	0,67	0,73	1,65	1,32	1,72	2,14	19
Total anthocyanes (mg/L)	0,82	0,76	0,68	25,21	31,75	2,19	1,74	16
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,85	0,80	0,36	0,22	0,39	2,33	1,33	17
SPP (Absorbance)	0,87	0,82	0,78	0,25	0,25	2,28	2,30	16
LPP (Absorbance)	0,91	0,88	0,67	0,34	0,33	2,60	2,64	14
PP (Absorbance)	0,91	0,89	0,83	0,46	0,33	2,72	3,86	13
Anthocyanes libres (mg/L)	0,69	0,62	0,62	67,51	71,54	1,71	1,61	11
<b>Tanins</b>								
BSA (mg/L)	0,92	0,90	0,86	70,63	65,07	2,93	3,18	14
MCP (mg/L)	0,77	0,70	0,53	392,42	317,18	1,64	2,03	13
Tanins totaux (mg/L)	0,86	0,82	0,81	661,25	507,37	2,17	2,83	11

Tableau 33 : Prédiction des anthocyanes et des tanins sur les vins commerciaux par les modèles IRTF, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Composition en flavan-3-ols</b>								
Flavan-3-ol (mmol/L)	0,86	0,79	0,78	2,55	2,24	2,08	2,37	10
Flavan-3-ol terminal (mmol/L)	0,66	0,54	0,52	1,63	1,86	1,44	1,26	10
Catéchine (mmol/L)	0,72	0,61	0,50	0,72	0,61	1,47	1,73	10
Epicatéchine (mmol/L)	0,92	0,87	0,77	1,11	1,09	2,62	2,67	11
Epigallocatechine (mmol/L)	0,82	0,61	0,60	1,49	1,78	1,63	1,37	14
Epicatéchine-gallate (mmol/L)	0,89	0,83	0,61	0,08	0,11	2,33	1,69	11
<b>Anthocyanes moléculaires</b>								
Total glucoside (mg/L)	0,88	0,85	0,74	16,71	19,82	2,52	2,12	10
Total acétyl (mg/L)	0,70	0,66	0,64	6,76	6,45	1,70	1,79	7
Total coumaroyl (mg/L)	0,72	0,69	0,57	1,68	1,75	1,68	1,62	6
Total anthocyanes (mg/L)	0,89	0,86	0,72	21,56	27,13	2,56	2,04	10
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,90	0,85	0,67	0,21	0,26	2,51	1,99	5
SPP (Absorbance)	0,82	0,81	0,52	0,27	0,27	2,12	2,12	11
LPP (Absorbance)	0,93	0,92	0,75	0,27	0,27	3,29	3,22	7
PP (Absorbance)	0,97	0,96	0,90	0,28	0,25	4,42	4,98	9
Anthocyanes libres (mg/L)	0,85	0,83	0,75	49,68	49,85	2,32	2,31	8
<b>Tanins</b>								
BSA (mg/L)	0,91	0,88	0,84	74,11	70,19	2,79	2,95	10
MCP (mg/L)	0,87	0,79	0,76	316,66	251,50	2,03	2,56	10
Tanins totaux (mg/L)	0,86	0,79	0,53	722,22	716,16	1,99	2,01	11

Tableau 34 : Prédiction des anthocyanes et des tanins sur les vins commerciaux par les modèles UV-visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Les paramètres décrivant la composition en flavan-3-ols montrent de bons résultats pour la concentration en unités flavan-3-ols totales et pour la concentration en épicatechine avec l'IRTF et l'UV-visible. Pour l'IRTF, nous obtenons un RPD CV de 1,84 et un RPD P de 2,02, et pour l'UV-visible nous obtenons un RPD CV de 2,24 et un RPD P de 2,37. Contrairement au modèle construit pendant la fermentation, les spectres UV-visible apparaissent plus performants que les spectres IRTF. Nous pouvons attribuer cette différence de robustesse à la forte variabilité de notre jeu de données qui affectera plus fortement les spectres IRTF où l'absorbance d'un plus grand nombre de composés est observable et pourra perturber la lecture. Les autres paramètres décrivant la composition en flavan-3-ols n'apparaissent pas assez robustes pour être envisagés (RPD P < 2). Contrairement aux résultats

de la fermentation pour les spectres IRTF, le jeu de données étudié sur les vins finis ne permet pas de prédire la concentration en épicatechine-gallate, et donc de prédire les teneurs en tanins de pépins et de pellicules.

Si l'on s'intéresse aux paramètres des anthocyanes moléculaires, nous constatons des difficultés à prédire ces derniers avec l'IRTF. Néanmoins les résultats sont encourageants. Pour la somme des anthocyanes glucosides nous obtenons un RPD CV de 2,11 et un RPD P de 1,67, pour la somme des anthocyanes acétylglucosides nous obtenons un RPD CV de 2,13 et un RPD P de 1,61, pour la somme des anthocyanes coumaroylglucosides nous obtenons un RPD CV de 1,72 et un RPD P de 2,14 et pour la somme des anthocyanes totales nous obtenons un RPD CV de 2,19 et un RPD P de 1,74. Ces résultats sont inférieurs à ceux obtenus pour le suivi de la fermentation.

Concernant les modèles construits avec les spectres UV-visible, nous constatons de meilleurs résultats pour la somme des anthocyanes glucosides (RPD CV : 2,52 ; RPD P : 2,12) et pour la somme des anthocyanes totales (RPD CV : 2,56 ; RPD P : 2,04). Les anthocyanes glucosides représentant plus de 50 % des anthocyanes, il est possible que les résultats de ce modèle soient dus à une corrélation indirecte. Néanmoins, les données obtenues montrent que pour les vins finis commerciaux, l'UV-visible apparaît plus adapté que l'IRTF dans la prédiction des anthocyanes totales. Cette observation est confirmée par les paramètres de l'évolution des anthocyanes, où l'UV-visible montre les meilleures robustesses. Nous constatons des résultats prédictifs aptes à être utilisés pour tous les paramètres étudiés, avec un modèle très robuste pour les pigments polymérisés (RPD P : 4,98). L'IRTF apparaît bien moins précis, mais indique néanmoins de bons résultats pour les petits pigments polymérisés (RPD P : 2,30), les grands pigments polymérisés (RPD P : 2,64) et la somme des pigments polymérisés (RPD P : 3,86).

Globalement, la comparaison des modèles obtenus avec les spectres UV-vis et IRTF pour les anthocyanes montre une capacité supérieure des spectres UV-vis à obtenir une bonne prédiction. La propension des anthocyanes à absorber dans le visible, et leur capacité à impacter la couleur en fonction de leur degré de polymérisation peuvent expliquer cette différence. De plus, les résultats obtenus avec les spectres IRTF apparaissent plus faibles que ceux de l'analyse de la fermentation. Cela peut s'expliquer par la grande variabilité des vins utilisés, avec une absorbance due aux nombreux composés organiques qui peuvent impacter la lecture dans l'IRTF. De plus, les concentrations plus faibles en anthocyanes moléculaires dans les vins commerciaux que dans les

échantillons suivis pendant la fermentation apportent une difficulté supplémentaire pour obtenir une prédiction efficace.

En s'intéressant à l'analyse des tanins condensés, on constate que la tendance précédemment constatée s'inverse. Pour deux des paramètres étudiés, les tanins précipitables à l'albumine de sérum bovin et les tanins totaux, les résultats pour les spectres IRTF sont plus robustes que ceux de l'UV-visible. Avec un RPD P pour les tanins précipitables à l'albumine de sérum bovin et les tanins totaux respectivement de 3,18 et 2,83, l'IRTF confirme son fort potentiel dans la prédiction des tanins, malgré une forte variabilité induite par les cépages et les millésimes. Les résultats des spectres UV-visible sont inférieurs, avec un RPD P pour les tanins précipitables à l'albumine de sérum bovin et les tanins totaux respectivement de 2,95 et 2,01, mais ils apparaissent également être une méthode fiable pour l'analyse des tanins condensés.

Contrairement aux deux paramètres présentés ci-dessus, les tanins précipitables à la méthylcellulose montrent des résultats très hétérogènes sur les deux zones spectrales étudiées. Alors que la prédiction avec les spectres UV-visible reste cohérente avec les résultats précédents, avec une valeur de RPD CV de 2,03 et de RPD P de 2,56, l'utilisation des spectres IRTF montre une baisse de la prédiction, avec une valeur de RPD CV de 1,64 et de RPD P de 2,03. Nous pouvons exclure le biais analytique car les résultats de l'UV-visible restent prédictifs. Il faut donc chercher l'explication de cette différence dans l'analyse elle-même. Le polymère de méthylcellulose interagit sans discernement avec les tanins condensés et les pigments polymérisés, et la différence de lecture à 280 nm ne les sépare pas, contrairement à l'analyse des tanins précipitables à l'albumine de sérum bovin. On peut donc émettre l'hypothèse que la diversité des composés précipités est trop importante pour les spectres IRTF pour l'étude des vins finis commerciaux.

### ***b) Combinaison des spectres IRTF et UV-visible pour la prédiction des anthocyanes et des tanins sur les vins finis commerciaux***

La combinaison des spectres IRTF et UV-visible a pu montrer précédemment des résultats encourageants, notamment sur l'étude du raisin pendant la maturité et sur l'étude du vieillissement. Nous avons donc décidé d'étudier cette combinaison pour les vins finis commerciaux. Les résultats obtenus sont présentés dans le Tableau 35.

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Composition en flavan-3-ols</b>								
Flavan-3-ol (mmol/L)	0,73	0,62	0,37	1,74	2,12	3,05	2,50	12
Flavan-3-ol terminal (mmol/L)	0,91	0,86	0,61	2,16	2,62	1,09	0,90	12
Catéchine (mmol/L)	0,85	0,76	0,16	0,57	0,75	1,85	1,41	13
Epicatéchine (mmol/L)	0,87	0,85	0,67	1,22	1,08	2,39	2,70	8
Epigallocatechine (mmol/L)	0,85	0,68	0,56	1,39	1,80	1,75	1,35	17
Epicatéchine-gallate (mmol/L)	0,77	0,74	0,61	0,09	0,11	2,07	1,69	7
<b>Anthocyanes moléculaires</b>								
Total glucoside (mg/L)	0,82	0,79	0,76	19,17	20,08	2,20	2,10	7
Total acétyl (mg/L)	0,79	0,76	0,73	5,59	5,93	2,06	1,94	8
Total coumaroyl (mg/L)	0,77	0,73	0,65	1,52	1,69	1,86	1,67	7
Total anthocyanes (mg/L)	0,82	0,80	0,76	24,43	26,88	2,26	2,05	7
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,88	0,85	0,82	0,21	0,18	2,41	2,86	11
SPP (Absorbance)	0,88	0,84	0,52	0,25	0,24	2,31	2,40	9
LPP (Absorbance)	0,95	0,93	0,78	0,25	0,23	3,54	3,76	10
PP (Absorbance)	0,98	0,96	0,84	0,26	0,24	4,83	5,26	15
Anthocyanes libres (mg/L)	0,92	0,90	0,79	36,90	45,90	3,12	2,51	11
<b>Tanins</b>								
BSA (mg/L)	0,96	0,92	0,86	60,95	61,64	3,39	3,36	17
MCP (mg/L)	0,88	0,80	0,73	315,23	258,68	2,04	2,49	13
Tanins totaux (mg/L)	0,91	0,87	0,42	558,71	891,91	2,57	1,61	12

Tableau 35 : Prédiction des anthocyanes et des tanins sur les vins par les modèles IRTF-UV-visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

En comparant ces résultats avec les résultats obtenus pour les spectres IRTF ou UV-visible, nous constatons une amélioration des prédictions pour les paramètres suivants : concentration en unités flavan-3-ols totales (RPD P : 2,50), concentration en unités épicatechines (RPD P : 2,70), concentration en anthocyanes acétylglucosides (RPD P : 1,94), concentration en anthocyanes moléculaires totales (RPD P : 2,05), teneurs en pigments monomériques (RPD P : 2,86), en petits pigments polymérisés (RPD P : 2,40), en grands pigments polymérisés (RPD P : 3,76), somme des pigments polymériques (RPD P : 5,26), concentration en anthocyanes libres (RPD P : 2,51) et tanins précipitables à la méthylcellulose (RPD P : 3,36). Nous constatons cependant que ces résultats restent

relativement proches du meilleur modèle obtenu par l'analyse IRTF ou UV-visible. Il est difficile d'attribuer cette augmentation de robustesse à une complémentarité des deux spectres. L'hypothèse la plus probable est que l'augmentation est due à une diminution du bruit assimilé à d'autres composés du vin, et permet ainsi une lecture plus précise du paramètre recherché. Globalement, le gain de prédiction sera négligeable par rapport aux instruments demandés, et nous préférons une analyse IRTF ou UV-visible seule.

Nous constatons que tous les paramètres relatifs aux anthocyanes indiquent une augmentation de la prédiction grâce à la combinaison des deux spectres. Il est alors possible d'envisager une complémentarité spécifique à ces composés. Considérant le rôle prépondérant du visible dans le dosage des anthocyanes, il est possible d'intégrer des longueurs d'onde spécifiques pour augmenter la robustesse des modèles IRTF. Les longueurs d'onde suivantes ont été sélectionnées : 420 nm, 520 nm et 620 nm. Le choix s'est porté sur ces valeurs car elles sont déjà utilisées comme paramètres de couleur dans l'industrie du vin, pour calculer l'intensité de la couleur et l'intensité de coloration modifiée. De plus, certains appareils IRTF, tel que le Winescan Flex avec le module couleur, permettent la lecture de ces longueurs d'onde. Ainsi, les modèles seraient directement applicables pour ces dispositifs si cette combinaison augmente la robustesse de la prédiction. Les résultats obtenus pour les paramètres des anthocyanes avec l'addition à l'IRTF des longueurs d'ondes du visible 420 nm, 520 nm et 620 nm sont présentés dans le Tableau 36.

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Anthocyanes moléculaires</b>								
Total glucoside (mg/L)	0,64	0,59	0,65	26,00	26,71	1,62	1,58	8
Total acétyl (mg/L)	0,80	0,75	0,70	5,26	7,43	2,19	1,55	12
Total coumaroyl (mg/L)	0,64	0,59	0,54	1,84	1,74	1,54	1,63	8
Total anthocyanes (mg/L)	0,80	0,74	0,71	26,89	32,07	2,05	1,72	15
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,77	0,74	0,53	0,27	0,32	1,90	1,61	7
SPP (Absorbance)	0,80	0,76	0,64	0,30	0,30	1,94	1,92	9
LPP (Absorbance)	0,94	0,91	0,73	0,28	0,28	3,10	3,20	16
PP (Absorbance)	0,96	0,94	0,91	0,34	0,30	3,70	4,19	15
Anthocyanes libres (mg/L)	0,73	0,70	0,70	63,23	60,39	1,82	1,91	8

Tableau 36 : Prédiction des anthocyanes sur les vins par les modèles IRTF combinés aux longueurs d'ondes 420, 520 et 620 nm du visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Si l'on compare ces résultats avec ceux obtenus précédemment pour l'IRTF seul, nous n'observons pas d'augmentation significative pour les paramètres des anthocyanes moléculaires. Les longueurs d'onde ajoutées n'apparaissent pas suffisantes pour améliorer la robustesse du modèle, et nécessitent pour cela le spectre UV-visible entier. Pour l'évolution des anthocyanes en revanche, nous observons une amélioration de la prédiction pour pratiquement tous les paramètres, à l'exception des petits pigments polymérisés. Pour les pigments monomériques, les résultats avancés (RPD CV : 1,90 ; RPD P : 1,61) n'apparaissent pas encore suffisants, mais montrent un bon potentiel. Les données de prédictions obtenues pour les grands pigments polymérisés (RPD CV : 3,10 ; RPD P : 3,20) et la somme des pigments polymérisés (RPD CV : 3,70 ; RPD P : 4,19) sont très robustes, ils peuvent être directement utilisables. Les anthocyanes libres semblent également prédictibles (RPD CV : 1,82 ; RPD P : 1,91), et nécessiteraient l'ajout d'autres échantillons pour permettre une meilleure calibration.

Dans l'ensemble, nous ne constatons pas d'amélioration pour la prédiction des anthocyanes moléculaires avec l'ajout des longueurs d'onde 420 nm, 520 nm et 620 nm. Cette observation montre bien qu'il faut faire la distinction entre la composition moléculaire, qui apparaît mieux prédite grâce à l'IRTF, et la couleur qui en découle, mieux prédite avec l'UV-visible. Pour l'évolution des

anthocyanes, qui impactent fortement l'évolution de la couleur, l'addition des longueurs d'onde du visible semblent tout à fait adaptée pour améliorer la robustesse et permettre à l'IRTF de modéliser ces paramètres. L'IRTF seul ne montrait pas de résultats suffisants pour cette évolution, l'ajout des paramètres d'intensité colorante pourrait donc permettre de développer des modèles avec une erreur de prédiction acceptable.

### **3. Modélisation de l'astringence et l'amertume par IRTF et UV-visible couplés à l'analyse PLS**

La littérature montre une corrélation forte entre les tanins précipitables à l'albumine de sérum bovin et l'astringence (325). Cependant, la seule étude sur le sujet porte seulement sur 21 vins du sud de la France, et le manque de diversité couplé à la collection réduite d'échantillons ne permet pas d'affirmer avec exactitude ce lien. De plus, nous savons que plusieurs paramètres peuvent influencer l'astringence, tels que la composition en unités flavan-3-ols, le degré de polymérisation moyen, ou l'effet matrice avec le pH, l'éthanol ou l'acidité (3,246,272). Concernant l'amertume, il n'existe pas à notre connaissance de travail similaire.

Nous avons cherché les corrélations possibles entre l'astringence et l'amertume et les analyses des polyphénols des vins. Sur les 92 vins étudiés, 42 ont été dégustés par un panel entraîné pour déterminer un indice d'astringence compris entre 0 et 10. Les vins sont originaires de 13 pays différents et composés de 29 cépages différents et de 15 millésimes différents (entre 1997 et 2017) afin d'obtenir la plus grande variabilité possible. Une analyse en composantes principales a été construite avec la majorité des paramètres étudiés (Figure 89). Les paramètres d'astringence et d'amertume ont été ajoutés en tant que variables externes.

Pour l'astringence, nous observons une très forte corrélation avec les tanins précipitables à l'albumine de sérum bovin, ce qui corrobore les résultats de la littérature (325). Ainsi, il est envisageable d'utiliser les valeurs de concentration des tanins précipitables à l'albumine de sérum bovin comme un indicateur fiable de la perception d'astringence.

Pour la perception d'amertume en revanche, nous ne constatons pas de corrélation avec les paramètres étudiés. Nous observons que sa perception semble légèrement corrélée avec la concentration en unités catéchines, cependant elle reste faible et ne permet pas d'en déduire un lien de causalité.

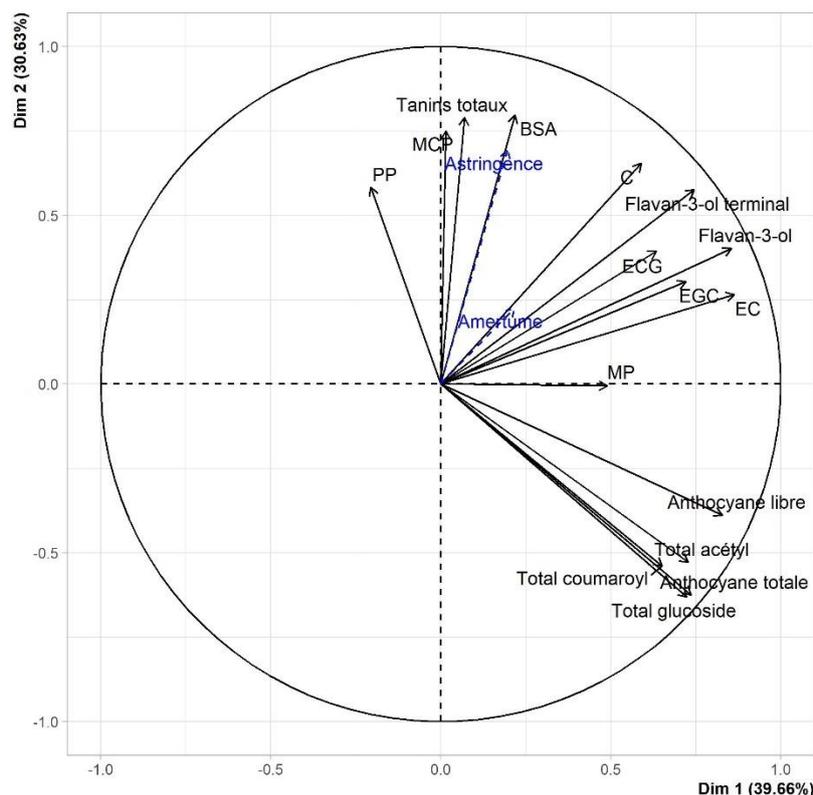


Figure 89 : Analyse en composantes principales des variables des vins dégustés, avec astringence et amertume en variables externes (MP : pigments monomériques, SPP : petits pigments polymériques, LPP : grands pigments polymériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, C : catéchine, EC : épicatechine, EGC : épigallocatechine, ECG : épicatechine-gallate)

L'astringence apparaît très fortement corrélée à la concentration en tanins précipitables à l'albumine de sérum bovin. Nous avons voulu approfondir cette observation en calculant les coefficients de régression linéaire des paramètres décrivant les tanins en fonction de l'astringence mesurée par le panel de dégustateurs (Tableau 37).

Paramètres	Coefficient de régression linéaire avec l'astringence
Flavan-3-ols totaux	0,22
Flavan-3-ols terminaux	0,28
Catéchine	0,27
Epicatechine	0,14
Epigallocatechine	0,17
Epicatechine-gallate	0,23
Tanins précipitables à la méthylcellulose	0,38
<b>Tanins précipitables à l'albumine de sérum bovin</b>	<b>0,64</b>
Tanins totaux	0,25

Tableau 37 : Coefficients de régression linéaire des paramètres décrivant les tanins en fonction de l'astringence mesurée par le panel de dégustateurs

Les résultats obtenus sont en accord avec l'ACP présenté sur la Figure 89. Nous observons une bonne corrélation de l'astringence avec les tanins précipitable à l'albumine de sérum bovin, très fortement supérieur aux corrélations avec les autres paramètres décrivant les tanins. Il apparaît donc possible d'estimer la perception d'astringence en modélisant cette concentration. Nous avons pu démontrer sur l'étude de la maturité du raisin, de la fermentation et des vins finis que ce paramètre pouvait être prédit par l'IRTF et l'UV-visible. Ces modèles peuvent donc être utilisés comme un indicateur fiable de l'astringence. A contrario, la perception d'amertume n'apparaît pas corrélée avec les paramètres étudiés décrivant la concentration et la composition en polyphénols. Ce travail préliminaire montre donc la difficulté à prédire cette perception.

#### ***a) Prédiction de l'astringence et de l'amertume par IRTF couplé à l'analyse PLS***

Afin d'évaluer le potentiel de l'IRTF à prédire les paramètres d'astringence et d'amertume des vins, nous avons cherché à modéliser directement l'indice de perception obtenu par la dégustation avec les données spectrales. Les modèles obtenus sont présentés sur la Figure 90. Le nombre d'échantillons étudiés n'étant pas suffisant pour utiliser une validation externe, nous avons utilisé la validation croisée.

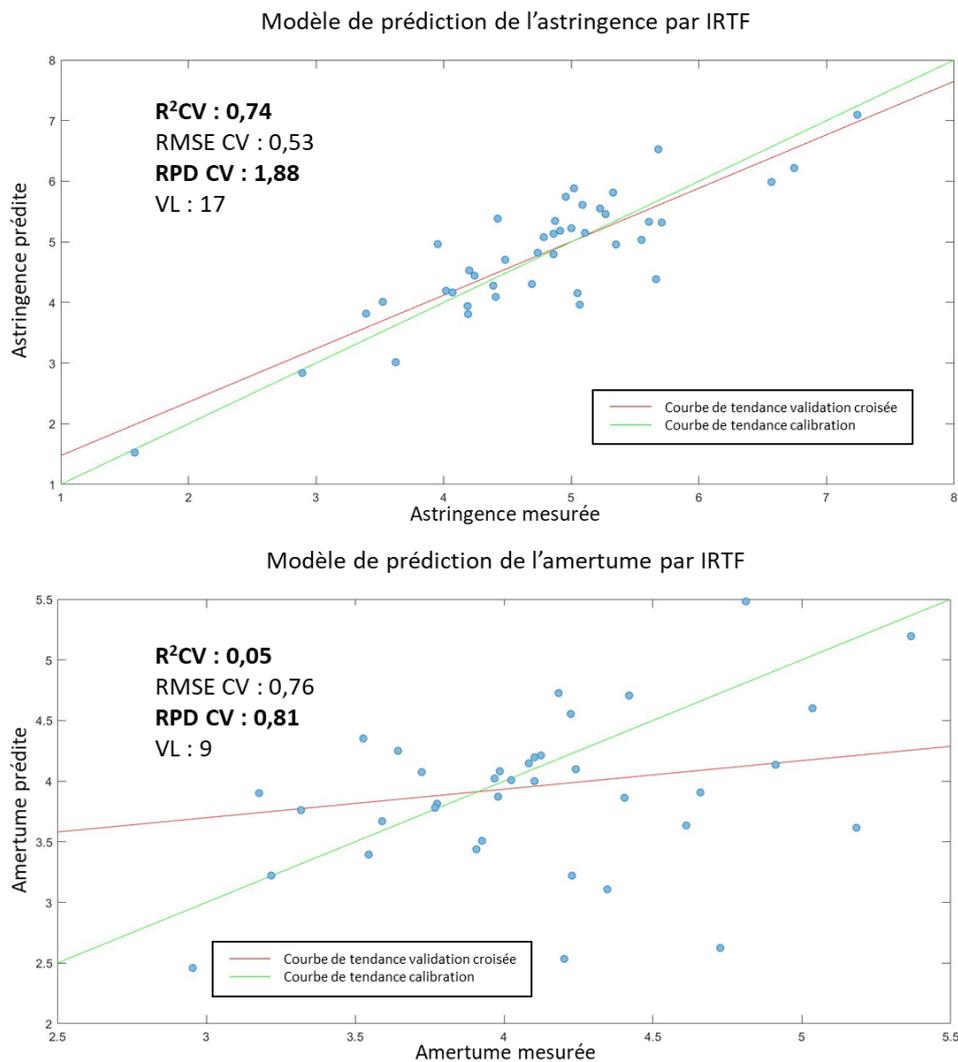


Figure 90 : Modèles de prédiction de l'astringence et de l'amertume mesurées par le panel de dégustateurs par IRTF couplé à l'analyse PLS ( $R^2_{CV}$  : coefficient de régression de validation croisée, RMSE CV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Pour l'astringence, nous observons des résultats avec un fort potentiel. Nous obtenons un  $R^2_{CV}$  de 0,74 qui montre une bonne robustesse de modélisation, et un RPD CV de 1,88 qui indique une erreur de prédiction acceptable. Nous avons ainsi mis en évidence pour la première fois qu'il était possible, grâce à l'IRTF couplée à l'analyse PLS, de modéliser une sensation gustative comme l'astringence afin de développer des modèles de prédiction. Le modèle doit désormais être approfondi avec l'ajout d'un plus grand nombre d'échantillons, afin d'être susceptible d'être utilisé. Cependant, si l'on considère les très bons résultats de prédiction que nous avons obtenus pour les tanins précipitables à l'albumine de sérum bovin et la forte corrélation de ce paramètre avec l'astringence, nous pouvons d'ores et déjà utiliser ce paramètre comme un indicateur fiable de la sensation d'astringence.

Concernant l'amertume, nous avons obtenu un modèle qui montre l'inaptitude de l'IRTF à prédire cette perception. Le  $R^2_{CV}$  de 0,05 indique une modélisation inefficace, et le RPD CV de 0,81 une erreur de prédiction au-delà des seuils acceptables. Sur la base de notre jeu d'échantillons, il n'apparaît pas possible d'utiliser l'IRTF pour prédire l'amertume.

### ***b) Prédiction de l'astringence et de l'amertume par UV-visible couplé à l'analyse PLS***

Nous avons effectué le même travail de modélisation avec l'UV-visible pour l'astringence et l'amertume. Les modèles obtenus sont présentés sur la Figure 91.

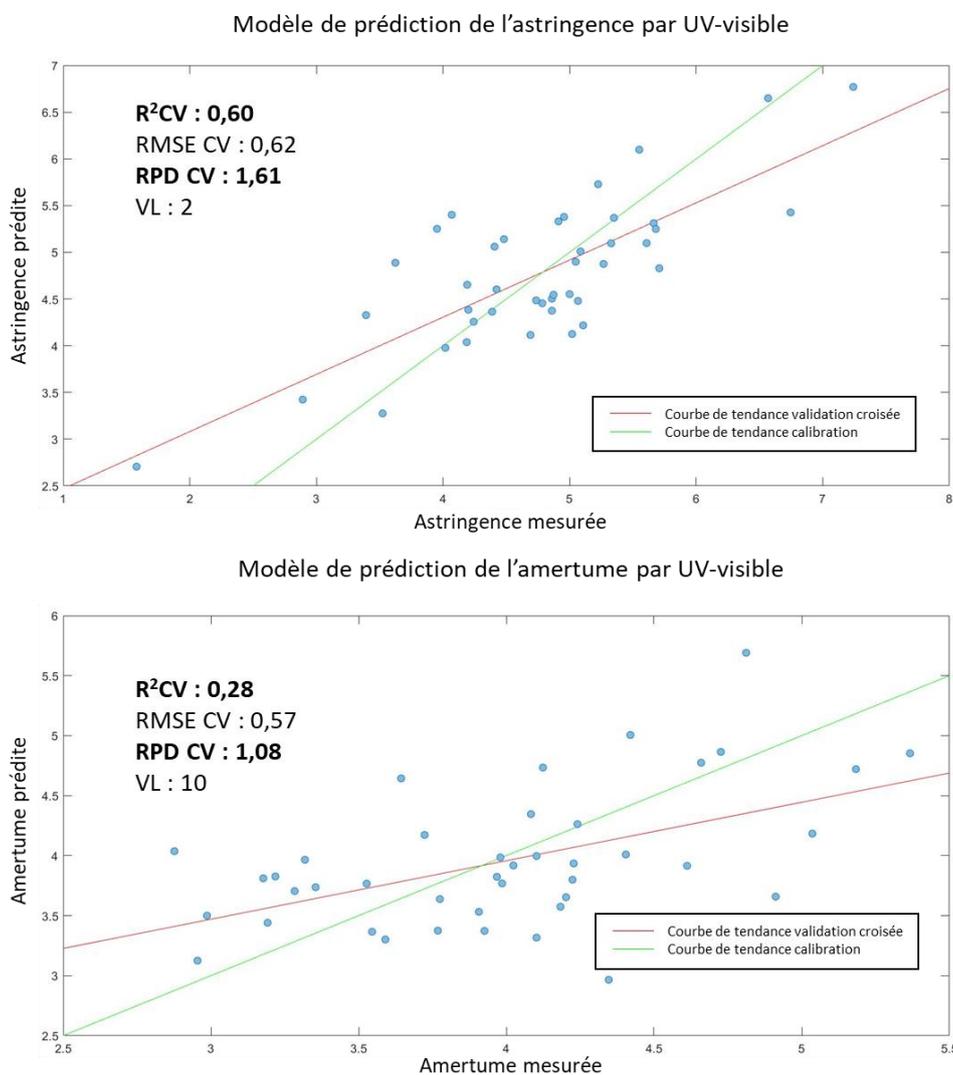


Figure 91 : Modèles de prédiction de l'astringence et l'amertume mesurées par le panel de dégustateurs par UV-visible couplé à l'analyse PLS ( $R^2_{CV}$  : coefficient de régression de validation croisée, RMSE CV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, VL : nombre de variables latentes)

Concernant l'astringence, nous observons une prédiction moins robuste qu'avec l'IRTF. Avec un  $R^2CV$  de 0,60 et un RPD CV de 1,61, la modélisation grâce à l'UV-visible apparaît potentiellement efficace, mais nécessite une augmentation de sa calibration et prédictivité pour être envisageable. Il lui sera préféré le modèle IRTF.

Pour l'amertume, les données recueillies sont supérieures à celle de l'IRTF. Néanmoins, avec un  $R^2CV$  de 0,28 et un RPD CV de 1,08, le modèle ne peut pas être considéré comme prédictif, ou possédant le potentiel de le devenir avec l'ajout d'un plus grand nombre d'échantillons. Au regard des données obtenues, nous pouvons avancer que ni l'UV-visible, ni l'IRTF ne permettent de prédire l'amertume. Cette perception apparaît trop complexe pour être modélisée grâce aux spectres étudiés, et ne montre pas de corrélation exploitable avec les analyses de références des polyphénols.

#### **4. Conclusions et perspectives de la prédiction des anthocyanes et des tanins par IRTF et UV-visible pendant l'élevage**

Grâce à cette étude portée sur la modélisation des concentrations en anthocyanes et en tanins sur les vins finis commerciaux, nous avons pu confronter la robustesse des prédictions face à une grande variabilité obtenue grâce à la diversité des vins en matière de cépages, de millésimes et d'origines. Les précédents chapitres nous ont permis de mettre en évidence la possibilité d'obtenir grâce à l'IRTF ou l'UV-visible une lecture rapide et fiable des concentrations en anthocyanes et en tanins pendant la maturité du raisin ainsi que pendant la vinification. Il apparaissait donc nécessaire de pouvoir continuer à analyser rapidement la composition phénolique du vin après sa mise en bouteilles et pendant son vieillissement, jusqu'à sa dégustation. La comparaison des deux spectres (IRTF et UV-visible) permet d'approfondir les avantages et les limites de chacun, afin d'utiliser la méthode la plus appropriée pour obtenir un résultat fiable.

Les résultats ont montré que dans l'ensemble, l'analyse des spectres UV-visible semblait plus appropriée dans la prédiction des polyphénols du vin, avec des résultats supérieurs au spectre IRTF dans l'analyse des anthocyanes, et des résultats très légèrement inférieurs dans l'analyse des tanins condensés. Cependant, les modèles IRTF montrent des résultats très concluants dans l'analyse des tanins condensés, et la polyvalence de cette méthode, qui peut également mesurer d'autres paramètres importants du vin comme les acides, l'éthanol ou les sucres, peut être préférée à la méthode UV-visible. Pour l'analyse des moûts et vins suivis pendant la vinification, l'IRTF a montré des résultats plus performants que ceux obtenus sur les vins finis commerciaux. Il apparaît donc que contrairement

à l'UV-visible, l'IRTF est plus sensible à une forte variabilité du jeu d'échantillons. Les prédictions obtenues montrent cependant qu'avec une calibration optimale composée d'une grande majorité de vin différents, nous pouvons obtenir des modèles de prédiction fiables qui s'affranchissent de la variabilité observée.

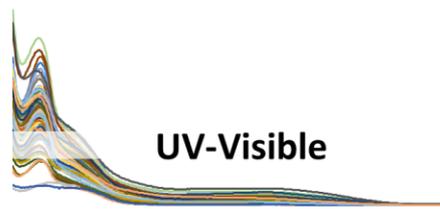
Afin de tester la complémentarité des deux zones spectrales, des modèles de prédiction ont été construits avec la combinaison des spectres IRTF et UV-visible. Si la plupart des paramètres étudiés montrent une légère augmentation de la prédiction, cela n'est pas suffisant pour indiquer une réelle complémentarité. Cependant, cette étude a montré la valeur du visible dans la prédiction des anthocyanes. Pour améliorer les modèles de prédiction IRTF pour ces composés, de nouveaux modèles ont été construits en ajoutant trois longueurs d'onde visibles spécifiques (420 nm, 520 nm et 620 nm) au spectre IRTF. Ces nouveaux modèles démontrent que l'ajout de ces longueurs d'onde peut compenser le manque de prédictivité des spectres IRTF dans le dosage des anthocyanes, et notamment dans l'estimation de la polymérisation de ces derniers.

La composition phénolique d'un vin joue un rôle prépondérant dans la perception d'astringence et d'amertume ressentie lors de sa dégustation. Si les concentrations mesurées peuvent être un bon indicateur de ces perceptions, elles ne sont parfois pas suffisantes pour les appréhender correctement. De plus, la littérature indique un nombre important de facteurs influençant l'astringence, tels que le degré de polymérisation moyen, la composition en flavan-3-ols ou la concentration en anthocyanes (3,246,273). Il apparaît alors difficile de discerner quels sont les paramètres qui influencent majoritairement cette sensation, et comment la modéliser correctement. Nous avons donc cherché à mettre en évidence les analyses de références associées à l'astringence et l'amertume. Les résultats obtenus montrent une forte corrélation de l'astringence avec la concentration en tanin précipitables à l'albumine de sérum bovin. Cette corrélation confirme une première constatation faite par une étude précédente sur un jeu d'échantillons moins important et moins varié (325). Il est alors possible d'utiliser la concentration en tanins précipitables comme un indicateur fiable de la perception d'astringence. Concernant l'amertume, le jeu de données étudié ne nous a pas permis d'identifier une possible corrélation avec les données de concentration et de composition des polyphénols recueillies.

Afin d'investiguer l'utilisation de l'IRTF ou de l'UV-visible dans la prédiction de l'astringence et de l'amertume, nous avons cherché à modéliser directement les valeurs de perception obtenues lors des dégustations. Pour l'astringence, nous avons obtenu de bons résultats avec l'IRTF. Cette méthode spectroscopique apparaît donc adaptée dans la prédiction de l'astringence des vins. Le modèle

construit à l'aide de l'UV-visible montre un léger manque de robustesse, et il lui sera préféré le modèle construit avec l'IRTF. Concernant l'amertume, aucune des deux méthodes spectroscopiques n'a montré de résultats suffisants.

Les modèles obtenus confirment l'utilité de l'IRTF et de l'UV-visible dans la prédiction de la composition phénolique des vins finis commerciaux. Un récapitulatif est présenté en Figure 92. Désormais, pour obtenir des modèles applicables au milieu industriel, il est nécessaire d'approfondir les différentes calibrations en y ajoutant un nombre plus important d'échantillons.



### Modélisations non exploitables

- Concentration en flavan-3-ol terminal
- Concentration en catéchine
- Concentration en épigallocatechine
- Concentration en épicatechine-gallate
- Amertume

- Concentration en flavan-3-ol terminal
- Concentration en catéchine
- Concentration en épigallocatechine
- Concentration en épicatechine-gallate
- Amertume

### Modélisations efficaces, nécessitent un développement approfondi

- Concentration en flavan-3-ol total
- Concentration en épicatechine
- Concentration en anthocyane glucoside
- Concentration en anthocyane acétylglucoside
- Concentration en anthocyane coumaroylglucoside
- Concentration en anthocyane totale
- Pigments monomériques (avec ajout visible)
- Concentration en anthocyanes libres
- Concentration en tanins précipitables à la méthylcellulose
- Astringence

- Concentration en flavan-3-ol total
- Concentration en épicatechine
- Concentration en anthocyane glucoside
- Concentration en anthocyane acétylglucoside
- Concentration en anthocyane coumaroylglucoside
- Tanins totaux
- Astringence

### Modélisations exploitables

- Pigments polymériques
- Concentration en tanins précipitables à l'albumine de sérum bovin
- Tanins totaux

- Concentration en anthocyane totale
- Pigments monomériques
- Pigments polymériques
- Concentration en anthocyanes libres
- Concentration en tanins précipitables à l'albumine de sérum bovin
- Concentration en tanins précipitables à la méthylcellulose

Figure 92 : récapitulatif de la modélisation des anthocyanes et des tanins sur les vins finis commerciaux par IRTF et UV-visible

## **E. Modélisation globale des anthocyanes et tanins des moûts et vins par IRTF et UV-visible couplé à une analyse de régression PLS**

Les études précédentes ont démontré l'aptitude de l'IRTF et de l'UV-visible couplé à une analyse de régression PLS pour prédire les anthocyanes et les tanins pendant la maturité du raisin, la vinification, et les vins commerciaux. La matrice étudiée a un impact important sur la qualité de la prédiction, c'est pourquoi nous avons étudié chaque partie de la vie du vin indépendamment. Nous pouvons désormais rechercher s'il est possible de concevoir une modélisation globale des anthocyanes et des tanins qui s'affranchisse de la variabilité observée dans chacun des chapitres précédents. Nous avons donc rassemblé l'ensemble des données des échantillons de la fermentation, de l'élevage en barriques et en bouteilles, ainsi que des vins commerciaux. Les données de la maturité du raisin ont été volontairement exclues. Considérant l'extraction utilisée pour obtenir l'échantillon analysable (macération à pH 3,2), la matrice apparaît trop différente des autres échantillons pour être combinée avec ceux-ci.

Nous avons décidé d'utiliser l'ensemble de ces échantillons pour prédire les paramètres qui avaient montré une modélisation efficace ou un fort potentiel dans les études précédentes, et qui pourraient apporter une information importante sur l'évolution, la concentration ou la composition phénolique. Nous avons sélectionné les paramètres suivants :

- La concentration en unités flavan-3-ols totales, qui permet d'obtenir une concentration indépendante de la réactivité des tanins et de leur réactivité.
- La concentration en epicatechine-gallate, qui, couplée à la concentration en unités flavan-3-ols totales, permet d'estimer les proportions en tanins de pépins et de pellicule.
- La concentration en anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales.
- Les teneurs en pigments monomériques et polymériques, qui permettent d'estimer l'évolution et la polymérisation des anthocyanes.
- La concentration en anthocyanes libres.
- La concentration en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose.

## 1. Modélisation globale des anthocyanes et tanins des moûts et vins par IRTF couplé à une analyse de régression PLS

Les résultats des modélisations avec les spectres IRTF sont présentés dans le Tableau 38. Pour confirmer la prédiction des modèles, nous avons utilisé la validation croisée, ainsi qu'une validation externe déterminée par l'algorithme de Kennard-Stone. Pour cette dernière, le jeu d'échantillons a été divisé en un set de calibration comprenant 66 % des échantillons, et un set de validation comprenant 34 % des échantillons.

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Composition en flavan-3-ols</b>								
Flavan-3-ol (mmol/L)	0,83	0,80	0,71	3,06	3,42	2,31	2,07	13
Epicatechine-gallate (mmol/L)	0,73	0,68	0,66	0,19	0,20	1,90	1,80	10
<b>Anthocyanes moléculaires</b>								
Total glucoside (mg/L)	0,91	0,85	0,80	43,00	37,57	2,37	2,71	18
Total acétyl (mg/L)	0,91	0,86	0,83	12,87	10,93	2,54	2,99	19
Total coumaroyl (mg/L)	0,90	0,84	0,74	5,13	5,30	2,41	2,33	17
Total anthocyanes (mg/L)	0,91	0,86	0,83	59,24	50,38	2,45	2,88	18
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,77	0,71	0,47	0,35	0,36	1,90	1,85	13
PP (Absorbance)	0,91	0,87	0,79	0,45	0,43	2,52	2,66	18
Anthocyanes libres (mg/L)	0,89	0,85	0,57	74,08	71,12	2,40	2,50	16
<b>Tanins</b>								
BSA (mg/L)	0,92	0,87	0,80	92,23	112,83	2,85	2,33	19
MCP (mg/L)	0,90	0,87	0,85	280,59	268,44	2,69	2,82	14

Tableau 38 : Prédiction des anthocyanes et des tanins sur les vins en fermentation, en élevage et commerciaux par les modèles IRTF, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Nous observons de très bons résultats (RPD CV et RPD P supérieur à 2) pour la majorité des paramètres étudiés. Seuls les modèles de la concentration en épicatechine-gallate (RPD CV : 1,90 ; RPD P : 1,80) et les teneurs en pigments monomériques (RPD CV : 1,90 ; RPD P : 1,85) montrent une robustesse légèrement inférieure à celle attendue pour être considéré comme exploitable. Pour les pigments, nous avons constaté lors de l'étude des vins commerciaux qu'il était possible d'augmenter la prédiction en ajoutant les longueurs d'onde du visible 420, 520 et 620 nm. Nous avons

donc construit un modèle avec ces longueurs d'ondes pour les pigments monomériques et polymériques. Les données de prédiction sont présentées dans le Tableau 39.

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,83	0,82	0,65	0,27	0,32	2,49	2,07	13
PP (Absorbance)	0,93	0,92	0,90	0,34	0,30	3,30	3,72	18

Tableau 39 : Prédiction des pigments monomériques et polymériques sur les vins en fermentation, en élevage et commerciaux par les modèles IRTF avec l'ajout des longueurs d'ondes du visible 420, 520 et 620 nm, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, PP : pigments polymériques, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Nous constatons que l'ajout des longueurs d'onde du visible permet d'augmenter suffisamment la production des modèles IRTF pour les rendre exploitables. Les appareillages IRTF possédant un module couleur capable de mesurer ces longueurs d'onde du visible sont donc en mesure de prédire avec une erreur de prédiction acceptable les pigments monomériques et polymériques, et ainsi permettre de suivre l'évolution de la polymérisation des anthocyanes.

Si l'on compare l'ensemble de ces résultats de prédictions avec ceux obtenus pour la fermentation ou les vins commerciaux seuls, nous constatons qu'ils apparaissent moins robustes que pour le suivi de la fermentation, mais plus prédictifs que pour les vins commerciaux. De manière générale, l'ajout des vins commerciaux semble apporter de l'incertitude dans la prédiction. Ainsi, pour prédire les paramètres étudiés lors du suivi de la fermentation, un modèle spécifique sera préféré. Pour le suivi de l'élevage et des vins commerciaux en revanche, les modèles globaux construits apparaissent plus prédictif qu'un modèle spécifique.

L'IRTF apparaît donc tout à fait adapté à la construction d'un modèle global fiable et prédictif pour tous les paramètres étudiés à l'exception de la concentration en épicatechine-gallate. Pour le suivi de la polymérisation des anthocyanes, il est nécessaire d'ajouter à la lecture les longueurs d'ondes 420, 520 et 620 nm.

## 2. Modélisation globale des anthocyanes et tanins des moûts et vins par UV-visible couplé à une analyse de régression PLS

Les résultats des modélisations avec les spectres UV-visible sont présentés dans le Tableau 40. Les conditions de modélisation et de validation utilisées sont identiques à celles utilisées pour l'IRTF.

Paramètre	R <sup>2</sup> C	R <sup>2</sup> CV	R <sup>2</sup> P	RMSECV	RMSEP	RPD CV	RPD P	VL
<b>Composition en flavan-3-ols</b>								
Flavan-3-ol (mmol/L)	0,84	0,78	0,84	3,24	2,94	2,18	2,41	17
Epicatechine-gallate (mmol/L)	0,83	0,72	0,79	0,19	0,16	1,90	2,25	18
<b>Anthocyanes moléculaires</b>								
Total glucoside (mg/L)	0,83	0,77	0,67	50,79	50,24	2,01	2,03	9
Total acétyl (mg/L)	0,80	0,70	0,69	19,01	15,34	1,72	2,13	9
Total coumaroyl (mg/L)	0,73	0,68	0,68	7,24	6,27	1,70	1,97	8
Total anthocyanes (mg/L)	0,83	0,76	0,70	74,63	68,15	1,94	2,13	9
<b>Evolution des anthocyanes</b>								
MP (Absorbance)	0,87	0,84	0,86	0,27	0,25	2,48	2,71	9
PP (Absorbance)	0,93	0,90	0,81	0,41	0,29	2,78	3,92	9
Anthocyanes libres (mg/L)	0,87	0,83	0,84	79,03	56,14	2,25	3,16	9
<b>Tanins</b>								
BSA (mg/L)	0,89	0,86	0,86	102,94	92,49	2,55	2,84	15
MCP (mg/L)	0,86	0,84	0,89	311,48	222,23	2,43	3,40	7

Tableau 40 : Prédiction des anthocyanes et des tanins sur les vins en fermentation, en élevage et commerciaux par les modèles UV-visible, validation croisée et validation externe déterminée par l'algorithme de Kennard-Stone (MP : pigments monomériques, PP : pigments polymériques, BSA : tanins précipitables à l'albumine de sérum bovin, MCP : tanins précipitables à la méthylcellulose, R<sup>2</sup>C : coefficient de régression de la calibration, RMSEC : erreur quadratique moyenne de calibration, R<sup>2</sup>CV : coefficient de régression de la validation croisée, RMSECV : erreur quadratique moyenne de validation croisée, RPD CV : écart prédictif résiduel de validation croisée, R<sup>2</sup>P : coefficient de régression de prédiction, RMSEP : erreur quadratique moyenne de prédiction, RPD P : écart prédictif résiduel de prédiction, VL : nombre de variables latentes)

Tout comme pour l'IRTF, nous observons de très bons résultats pour la majorité des paramètres étudiés. Cependant, contrairement à l'IRTF, le modèle des pigments monomériques apparaît directement exploitable (RPD CV : 2,48 ; RPD P : 2,71), et le modèle de la concentration en épicatechine-gallate présente un léger manque de fiabilité (RPD CV : 1,90 ; RPD P : 2,25), mais avec un fort potentiel. Les modèles des anthocyanes glucosides (RPD CV : 2,01 ; RPD P : 2,03), acétylglucosides (RPD CV : 1,72 ; RPD P : 2,13), coumaroylglucosides (RPD CV : 1,70 ; RPD P : 1,97) et totales (RPD CV : 1,94 ; RPD P : 2,13) montrent une prédiction à la limite du seuil d'acceptabilité, ou très légèrement supérieure. Ces données confirment l'observation faite lors du

suivi de la fermentation, et montrent que l'IRTF est plus adapté à la modélisation des anthocyanes moléculaires.

En comparant la modélisation globale avec les modèles construits sur les vins en fermentation et commerciaux, nous obtenons les mêmes constatations que pour l'IRTF. La modélisation globale ne permet pas d'obtenir un modèle plus performant que le meilleur des modèles construits avec les vins en fermentation ou commerciaux.

De manière générale, l'utilisation d'un large panel d'échantillons représentant toutes les étapes de la vie d'un vin peut être utilisé pour construire des modèles efficaces afin de prédire les anthocyanes et les tanins. Ces modèles peuvent s'avérer extrêmement utiles, car ils permettent une lecture rapide et fiable quelle que soit la provenance de l'échantillon ou l'étape de la vinification. Cependant, la conception d'un modèle global n'arrive pas à s'affranchir totalement de la forte variabilité des différentes matrices (moût en fermentation, élevage, vin). Ainsi, pour obtenir la prédiction la plus précise possible, il apparaît nécessaire de développer des modèles spécifiques à chaque matrice.

---

## **Conclusion générale et perspectives**

---



Au cours de nos travaux, nous avons cherché à développer des modèles de prédiction efficaces des concentrations et compositions en anthocyanes et tanins du raisin, du moût et du vin rouge grâce à l'analyse spectrale IRTF et UV-visible couplée à l'analyse de régression multivariée PLS. Nous avons voulu mettre en évidence les avantages et les limites de chacune des analyses spectrales afin d'obtenir les meilleures modélisations possibles, qui puissent être fiables et prédictives pour un usage dans le milieu industriel viti-vinicole. Pour atteindre ces objectifs, nous avons séquencé nos recherches en 6 étapes clés (Figure 93) :

1. Echantillonnage des raisins, moûts et vins. La bonne sélection des échantillons utilisés pour la calibration de modèles de prédiction est primordiale. Ils doivent refléter la variabilité des raisins, moûts et vins qui seront ensuite analysés en routine. Nous avons donc recherché la plus grande diversité possible en termes de millésime, de cépage et de composition phénolique. Afin de s'assurer de la variabilité du jeu d'échantillons ainsi recueillis, nous avons étudié sa répartition phénolique à l'aide d'analyses en composantes principales ainsi que son évolution au cours du temps, qu'il s'agisse de la maturité, de la vinification ou du vieillissement.
2. Analyses spectrales. Il est nécessaire que l'enregistrement du spectre nécessaire à la prédiction soit le plus simple et rapide possible, afin d'être facilement applicable à l'analyse de routine. Hormis pour les échantillons de raisins, qui nécessitent une extraction préalable, nous avons travaillé directement sur les moûts et les vins. Pour l'UV-visible, nous avons utilisé une dilution au 1/100 qui rend directement exploitable la partie UV et visible sans saturation du spectre. Pour l'IRTF, aucune manipulation préliminaire n'a été nécessaire.
3. Analyses de références. Nous avons sélectionné des analyses décrites dans la littérature qui permettent d'étudier la composition et la concentration des tanins et anthocyanes. Nous avons pu déterminer sur nos échantillons leur composition et concentration en flavan-3-ols, leur composition et concentration en anthocyanes moléculaires, nous avons pu suivre les teneurs en pigments monomériques et polymérisés, ainsi que la concentration globale en tanins précipitables avec une protéine et un polysaccharide. Ces analyses nous ont permis de mesurer la concentration globale des tanins et anthocyanes et leur composition moléculaire. Elles nous ont aussi permis d'évaluer l'évolution et la polymérisation des anthocyanes, ainsi que la réactivité des tanins.

4. Analyses sensorielles. Des paramètres comme l'astringence et l'amertume sont importants pour le dégustateur. Nous avons cherché à modéliser ces perceptions à l'aide des analyses de références et des analyses spectrales. Pour cela, nous avons fait déguster une sélection de vins rouges à un panel de dégustateurs entraînés, afin de déterminer un indice d'astringence et d'amertume que nous avons cherché à prédire.
5. Création de modèles de prédiction. Nous avons utilisé la régression PLS sur les données spectrales et les analyses de références dans le but de concevoir des modèles capables de prédire les concentrations et compositions en tanins et anthocyanes. Le seul prétraitement appliqué aux données a été un centrage par la moyenne, pour ne pas dénaturer l'information. Nous avons aussi recherché un nombre de variables latentes le plus faible possible, afin d'éviter une mauvaise interprétation des résultats.
6. Validation des modèles de prédiction. Afin de s'assurer de la robustesse de nos modélisations, nous avons utilisé plusieurs méthodes de validation quand cela était possible : la validation croisée, la validation externe par millésime et la validation externe déterminée par l'algorithme de Kennard-Stone. Ces différentes méthodes nous ont permis de nous assurer de la fiabilité des modèles créés, et de rechercher les limites de prédiction face à un jeu d'échantillons externe à la calibration.

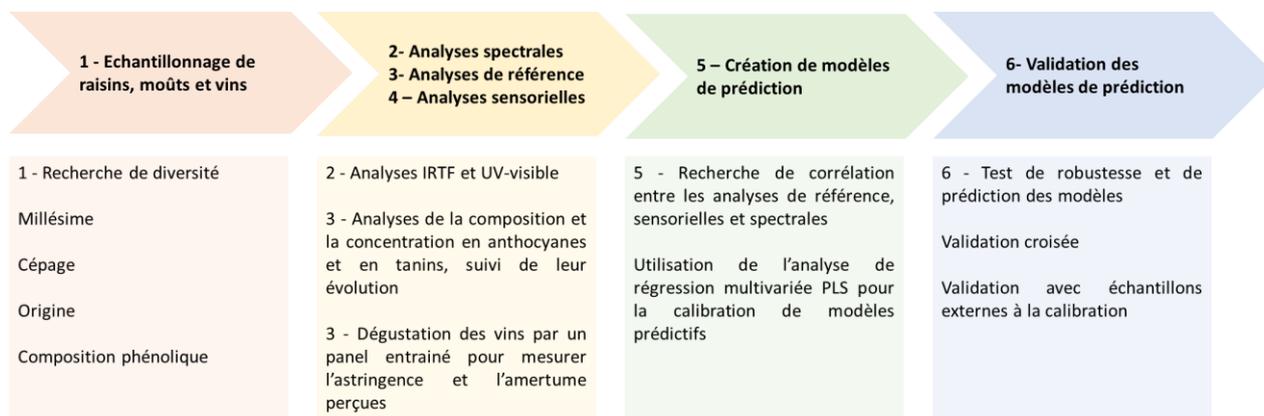


Figure 93 : Méthodologie de conception et validation de modèles de régression PLS grâce à l'IRTF et l'UV-visible pour la prédiction des concentrations et compositions en anthocyanes et tanins des raisins, moûts et vins.

Considérant le fort impact de la matrice du raisin, du moût ou du vin sur la réponse spectrale, nous avons divisé notre étude en fonction de la matrice étudiée. Dans un premier temps, nous avons évalué la variabilité et la dispersion des échantillons afin de s'assurer que le jeu de données obtenu était adapté à la conception de modèle de prédiction. Nous avons également pu étudier l'évolution des

paramètres étudiés en fonction du temps. Les données recueillies sur les échantillons de chaque matrice ont montré une variabilité suffisante pour concevoir des modèles de prédiction représentatifs. De plus, aucun échantillon divergent n'a été répertorié et écarté des analyses.

❖ Modèles de prédiction IRTF et UV-visible de la concentration en anthocyanes et en tanins pour le suivi de la maturité du raisin

Dans le cas du suivi de la maturité du raisin, les résultats obtenus ont mis en évidence le potentiel de l'IRTF et de l'UV-visible pour prédire les concentrations en anthocyanes et en tanins. Concernant l'IRTF, nous avons créé des modèles prédictifs (écart prédictif résiduel de validation croisée supérieur à 2) pour la majorité des paramètres, à savoir ApH1 et ApH3,2 de la méthode Glories, les concentrations en anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales et les concentrations en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose. Ces modélisations ont pu être validées par une validation externe déterminée par l'algorithme de Kennard-Stone. Cependant, quand nous avons cherché à approfondir cette validation en fonction des millésimes, nous avons constaté que la variabilité observée en 2018 n'était pas suffisante pour prédire le millésime 2019. Seul le modèle de la somme des anthocyanes moléculaires a montré des résultats de prédiction assez robustes. Il apparaît alors que ces paramètres peuvent être modélisés et prédits par l'IRTF, mais la calibration nécessite une forte variabilité des données pour être efficace. Dans le but de créer des modèles fiables et prédictifs, il est nécessaire d'ajouter de nouveaux échantillons avec différents millésimes, différents cépages, et différentes méthodes de viticulture.

Les modèles construits pour les résultats de l'UV-visible se sont avérés légèrement moins performants que ceux de l'IRTF. Nous avons été en mesure d'obtenir un écart prédictif résiduel de validation croisée (RPD CV) supérieur à 2 pour les paramètres suivants : les concentrations en anthocyanes glucosides, acétylglucosides et totales et les concentrations en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose. Tout comme pour l'IRTF, ces modèles ont pu être validés par une validation externe déterminée par l'algorithme de Kennard-Stone, mais ont démontré un manque de robustesse avec la validation en fonction du millésime. Sur le jeu de données étudié, il apparaît que les modèles construits avec l'UV-visible sont bien plus sensibles à l'effet millésime que les modèles construits avec l'IRTF.

Nous avons également recherché la possibilité de complémentarité des spectres IRTF et UV-visible en construisant des modèles combinant ces deux méthodes spectroscopiques. Si les résultats obtenus

ont montré des prédictions plus robustes pour la somme des anthocyanes moléculaires, cette hypothèse n'a pas pu être vérifiée avec les autres paramètres étudiés.

Les très bons résultats obtenus pour la prédiction de chacune des anthocyanes moléculaires, au regard de leur faible concentration, nous ont amené à questionner les limites de l'analyse multivariée appliquée au raisin. Nous avons pu démontrer qu'il existait une corrélation entre ces concentrations et la concentration de la somme des anthocyanes moléculaires. La prédiction de chacune des anthocyanes moléculaires distinctes a ainsi pu être attribuée à la modélisation de la somme des anthocyanes moléculaires, et non à l'empreinte spectrale du composé recherché. Il apparaît donc important, lors de la conception de modèle PLS à partir d'une matrice complexe, d'étudier en profondeur sa composition afin de s'assurer de ne pas mal interpréter les résultats obtenus.

L'IRTF et l'UV visible se sont montrés tout à fait adaptés pour suivre les concentrations en anthocyanes et en tanins pendant la maturité du raisin. Pour obtenir des modèles exploitables, il est désormais nécessaire de poursuivre cette étude avec l'ajout de nouveaux échantillons. De plus, il est important de travailler sur une méthode d'extraction des anthocyanes et des tanins efficace et rapide qui reste représentative de l'échantillon étudié, afin de réduire le temps d'analyse. La Figure 94 présente un récapitulatif des résultats obtenus pour le suivi de la maturité du raisin.

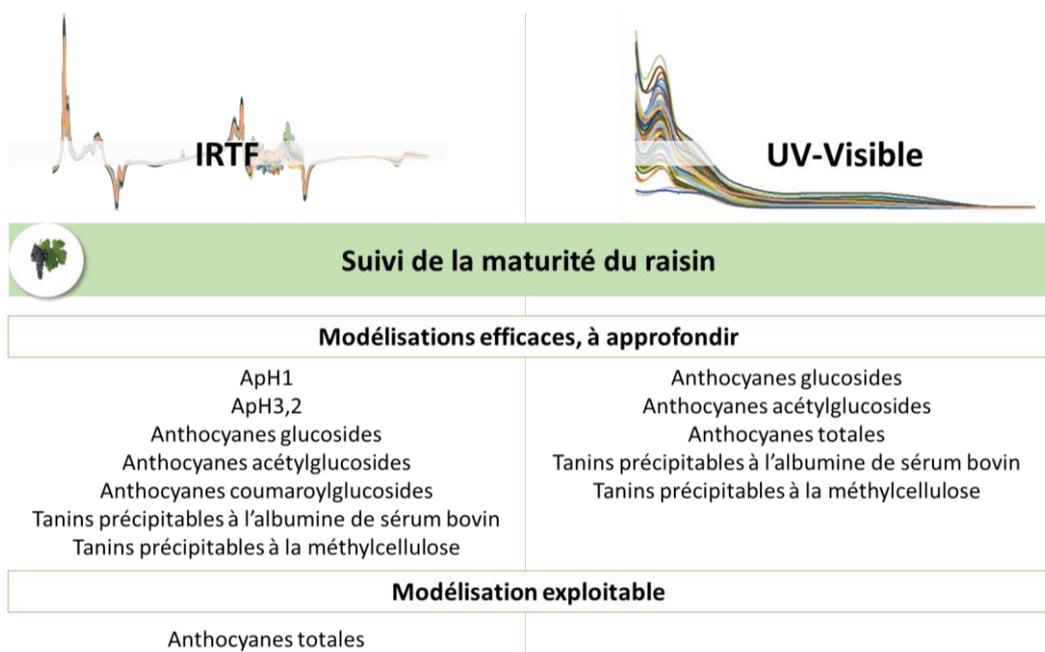


Figure 94 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour le suivi de la maturité du raisin

❖ Modèles de prédiction IRTF et UV-visible de la concentration en anthocyanes et en tanins pour le suivi de la fermentation des moûts et des vins

Pendant la fermentation des raisins et des moûts, nous nous sommes intéressés à la prédiction des concentrations en anthocyanes et en tanins. Les résultats de l'IRTF et de l'UV-visible se sont montrés concluants, et la majorité des paramètres étudiés ont été modélisés efficacement. Pour l'IRTF, nous avons été en mesure de produire des modèles prédictifs pour les concentrations en anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales, en pigments monomériques et polymériques, en anthocyanes libres et pour la concentration des différentes unités flavan-3-ols, à l'exception de l'épigallocatechine. Ces modèles doivent désormais être approfondis et validés avec l'ajout de nouveaux échantillons venant de cépages, millésimes et vinifications différents. Nous avons démontré que l'impact du millésime sur la prédiction était important et ne pouvait être négligé, comme pour le suivi de la maturité du raisin. Cependant, les modèles de prédiction des tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose ont démontré pouvoir s'affranchir de cet effet, et être directement exploitables. Nous avons également pu valider la surinterprétation des résultats des anthocyanes moléculaires distinctes soulevée pendant le suivi de la maturité du raisin.

Concernant l'UV-visible, tout comme pour le raisin, nous avons observé des résultats légèrement inférieurs à ceux obtenus par l'IRTF. Nous avons obtenu des modèles de prédiction avec un fort potentiel pour les concentrations des différentes unités flavan-3-ols, à l'exception de l'épigallocatechine, et les pigments monomériques et polymériques. De plus, les modèles des tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose et des anthocyanes libres se sont avérés assez robustes pour s'affranchir de l'effet millésime et être exploitables directement. Par comparaison des deux méthodes spectroscopiques, il apparaît que l'UV-visible est tout à fait adapté pour la prédiction des concentrations globales en tanins et anthocyanes, alors que l'IRTF, en plus de ces paramètres, permet d'approfondir la composition moléculaire. La Figure 95 présente un récapitulatif des résultats obtenus pour le suivi de la fermentation alcoolique des vins rouges.

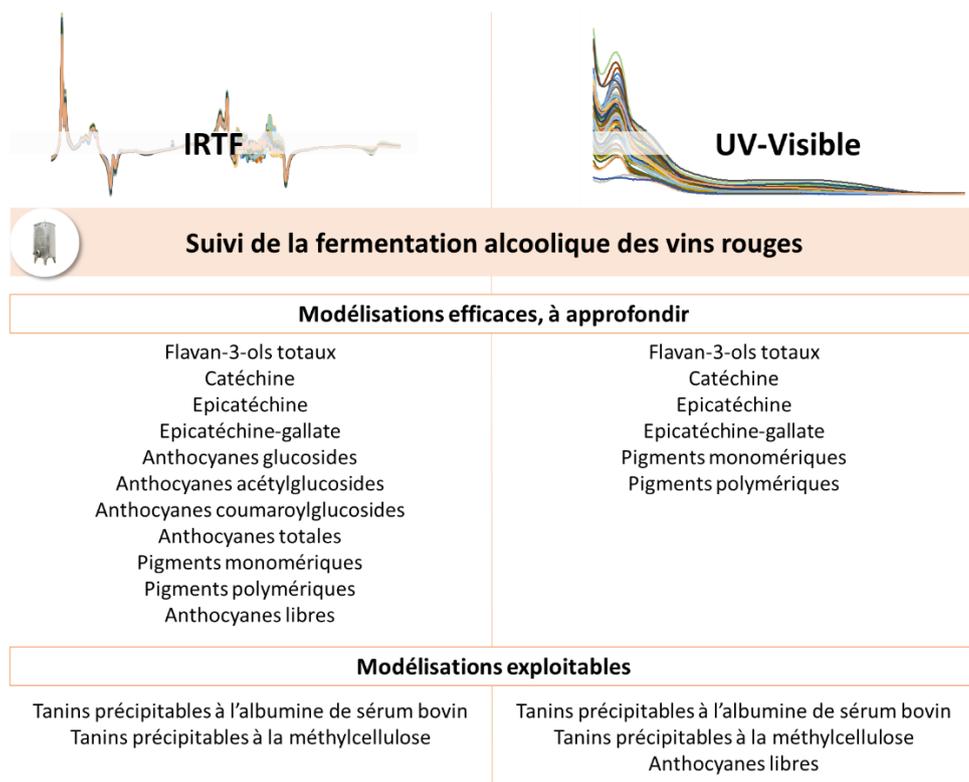


Figure 95 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour le suivi de la fermentation alcoolique des vins rouges

❖ Modèles de prédiction IRTF et UV-visible pour la discrimination de l'élevage en barriques et en bouteilles

L'élevage des vins en barriques et en bouteilles a été suivi et grâce à l'IRTF et l'UV-visible, nous avons été en mesure de discriminer les deux types d'élevage convenablement, et de montrer qu'il était possible en utilisant ces méthodes spectroscopiques de différencier le type de vieillissement. En utilisant la combinaison IRTF et UV-visible, nous avons été en mesure d'obtenir une discrimination parfaite. Afin de s'assurer du bon fonctionnement de ces modèles, il est désormais intéressant d'approfondir l'étude avec de nouveaux échantillons, en recherchant notamment l'impact du bois avec l'ajout de vins élevés avec des morceaux de bois de chênes. Nous avons également cherché à prédire les concentrations en anthocyanes et en tanins durant l'élevage grâce aux modèles construits sur les échantillons en fermentation. Malheureusement, les résultats obtenus ne se sont pas montrés assez robustes pour être exploitables. Les différences des vins en fermentation et en élevage sont trop importantes et impactent trop fortement les spectres IRTF et UV-visible, ne permettant pas une prédiction fiable. Il est alors nécessaire d'ajouter à la calibration un jeu d'échantillons provenant de l'élevage pour permettre une modélisation efficace.

❖ Modèles de prédiction IRTF et UV-visible de la concentration en anthocyanes et en tanins pour l'étude des vins finis commerciaux

Nous avons également travaillé sur des vins finis commerciaux en vue de prédire le contenu en anthocyanes et en tanins. Nous avons construit nos modèles grâce à un jeu d'échantillons cherchant à couvrir la plus grande variabilité possible en termes de cépages, de millésimes et d'origines. Nous avons obtenu pour l'IRTF des résultats satisfaisants pour une majorité de paramètres : les concentrations en unités flavan-3-ols totales et épicatechine, les concentrations en anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales, la concentration en anthocyanes libres et la concentration en tanins précipitables à la méthylcellulose. De plus, plusieurs modèles montrent des robustesses suffisantes pour être directement exploitables : la teneur en pigment polymérique, les concentrations en tanins précipitables à l'albumine de sérum bovin et en tanins totaux. Ces nouveaux modèles créés grâce à l'IRTF sont performants pour prédire les tanins, mais montrent des résultats inférieurs à ceux obtenus pendant la fermentation des moûts pour les concentrations en anthocyanes. Considérant la forte variabilité du jeu de données et les faibles concentrations en anthocyanes, il est nécessaire de continuer à développer la calibration de ces composés avec un apport de nouveaux échantillons.

Les résultats obtenus pour l'UV-visible se sont montrés plus performants que ceux de l'IRTF, notamment pour la prédiction des anthocyanes. L'UV-visible apparaît moins sensible à la forte variabilité du jeu de données, et donc plus adapté pour la prédiction des concentrations en tanins et anthocyanes qui composent les vins finis commerciaux. Si l'on considère les résultats obtenus sur le suivi de la maturité du raisin et de la fermentation, où nous avons observé des résultats supérieurs pour l'IRTF, il serait nécessaire de poursuivre l'étude des vins finis commerciaux. En effet, il est possible qu'avec un jeu d'échantillons plus important, il soit possible d'améliorer sensiblement la calibration de l'IRTF.

Les modèles de prédiction construits avec la combinaison de l'IRTF et de l'UV-visible ont montré des résultats légèrement supérieurs, mais pas suffisamment élevés pour conclure à une complémentarité des deux spectres. Cependant, ils nous ont permis de mettre en évidence la forte utilité du visible dans la prédiction de la composition et concentration en anthocyanes des vins finis commerciaux. Certains appareillages IRTF utilisés pour l'analyse du vin sont équipés d'un module couleur qui permet de lire des longueurs d'onde du visible spécifiques (420 nm, 520 nm et 620 nm). Nous avons donc cherché à améliorer les modèles de prédiction l'IRTF en y ajoutant ces longueurs

d'onde. Ces nouvelles modélisations ont démontré que l'on peut compenser le manque de prédictivité des spectres IRTF dans le dosage des anthocyanes et le suivi de leur polymérisation en ajoutant ces longueurs d'onde spécifiques du visible.

La possibilité de prédire les perceptions d'astringence et d'amertume a été étudiée grâce aux analyses de références ou avec l'analyse spectroscopique. Une sélection de vins finis commerciaux a été testée par un panel de dégustateurs entraînés afin de déterminer une valeur d'astringence et d'amertume pour chaque vin. Ces valeurs ont ensuite été confrontées aux données analytiques recueillies pour chercher de possibles corrélations. Nous avons ainsi démontré que la concentration en tanins précipitables à l'albumine de sérum bovin pouvait être utilisée comme un bon indicateur de la perception d'astringence, à l'instar des autres paramètres étudiés. De plus nous avons été en mesure de produire des modélisations démontrant d'un fort potentiel pour la prédiction de l'astringence grâce à l'IRTF et l'UV-visible. Cette nouvelle approche pourra être approfondie afin d'enrichir cette calibration pour obtenir les modèles les plus fiables possibles. Concernant l'amertume, cette sensation s'est avérée trop complexe à modéliser, et n'a pas montré de corrélation avec la composition en polyphénols. La Figure 94 présente un récapitulatif des résultats obtenus pour le suivi des vins finis commerciaux.

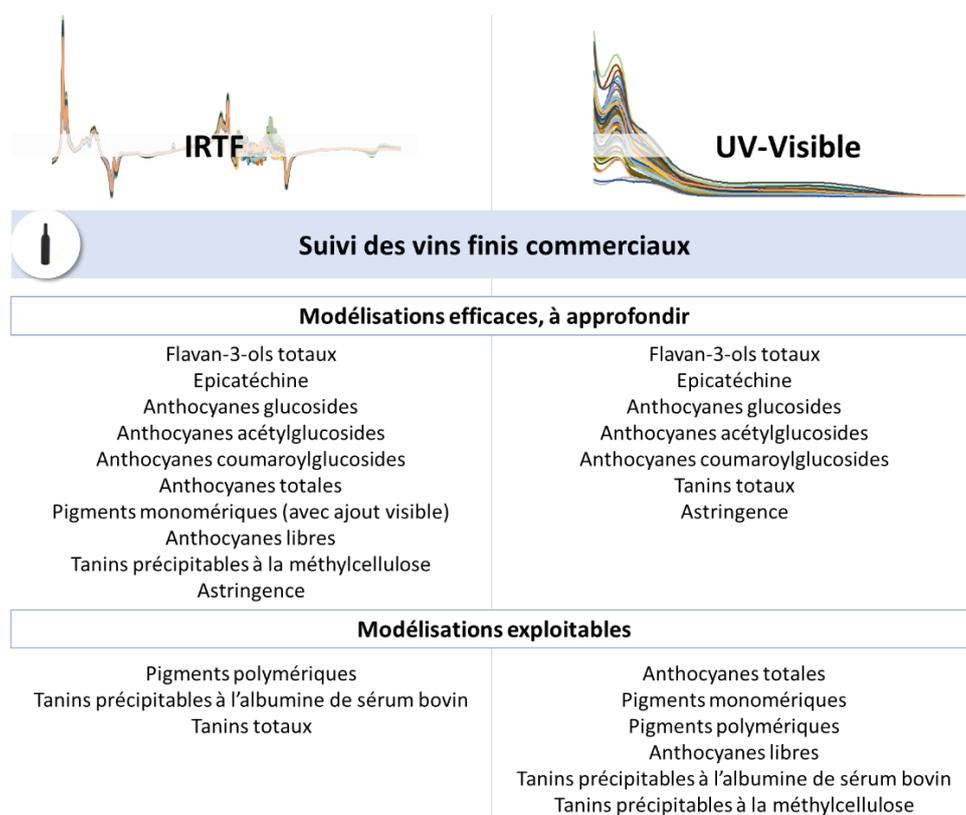


Figure 96 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour le suivi des vins finis commerciaux

❖ Modèles de prédiction globaux IRTF et UV-visible de la concentration en anthocyanes et en tanins pour le suivi de la fermentation, l'élevage et les vins commerciaux

L'étude du couplage des différentes matrices (raisin, moût et vin) nous ont démontré qu'il était possible en utilisant l'IRTF et l'UV-visible associés à l'analyse multivariée PLS de prédire efficacement les concentrations et compositions en anthocyanes et en tanins. Nous avons également pu mettre en évidence que la forte variabilité de la composition moléculaire des vins rouges en fonction du cépage, du millésime ou de la vinification pouvait fortement impacter la prédiction. La poursuite de ce travail a cherché s'il était possible de s'affranchir de cette variation, en construisant des modèles de prédiction grâce à l'IRTF et l'UV-visible pour la concentration et la composition en anthocyanes et en tanins utilisant les données recueillies pendant la vinification, l'élevage et sur les vins finis commerciaux. Nous avons obtenu un RPD P supérieur à 2 pour pratiquement tous les paramètres étudiés (concentration en flavan-3-ols et épicatechine-gallate, en anthocyanes glucosides, acétylglucosides, coumaroylglucosides et totales, en anthocyanes libres, en pigments monomériques et polymériques, en tanins précipitables à l'albumine de sérum bovin et à la méthylcellulose), à l'exception de la concentration en épicatechine-gallate et en pigments monomériques pour l'IRTF, et de la concentration en anthocyanes coumaroylglucosides pour l'UV-visible. Nous avons également démontré que l'ajout de longueurs d'onde du visible (420, 520 et 620 nm) à l'IRTF permettait d'améliorer significativement la prédiction des pigments monomériques et polymériques, pour suivre efficacement l'évolution de la polymérisation des anthocyanes.

Les modélisations construites à partir des échantillons de la vinification, de l'élevage et des vins finis commerciaux nous ont permis de prédire efficacement les concentrations en anthocyanes et en tanins sur toute la durée de vie d'un vin rouge, du premier jour d'encuvage jusqu'à sa dégustation. Ces modèles peuvent s'avérer très utiles pour le milieu industriel viti-vinicole car ils permettent d'obtenir rapidement l'empreinte polyphénolique d'un échantillon indépendamment de sa provenance ou de l'étape de vinification. Cependant, les modèles globaux ont démontré des robustesses inférieures aux modèles construits spécifiquement pour la fermentation ou les vins commerciaux. Si l'on recherche l'erreur de prédiction la plus faible possible, il est alors nécessaire de développer une modélisation spécifique pour la fermentation, l'élevage ou les vins commerciaux. La Figure 97 présente un récapitulatif des résultats obtenus pour la modélisation globale (fermentation, élevage et vins commerciaux).

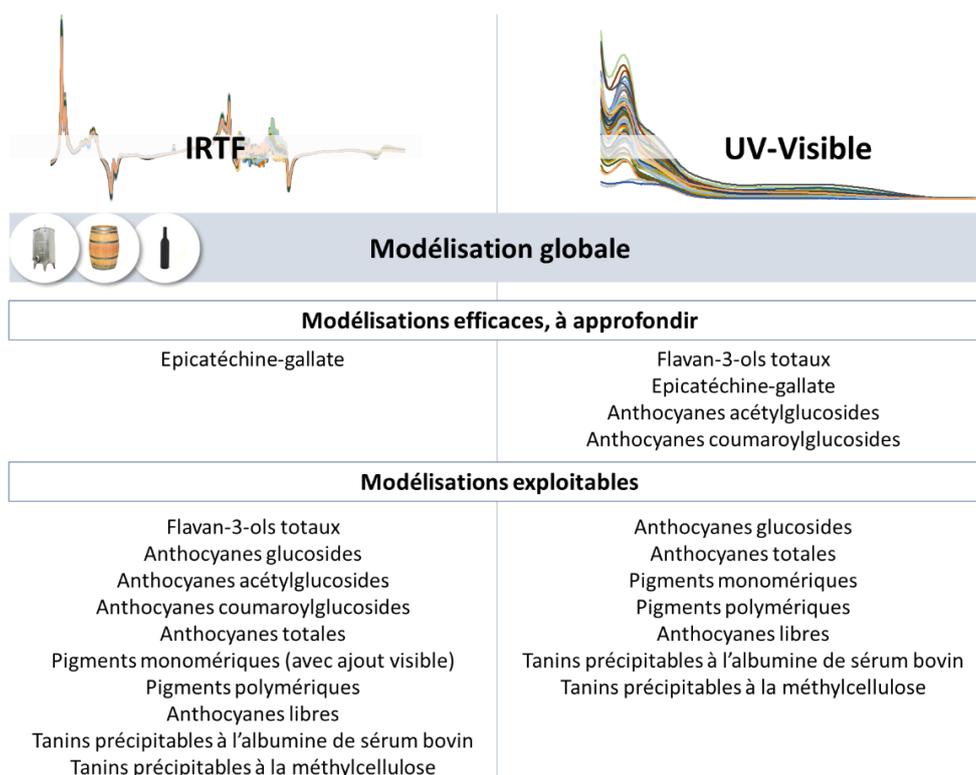


Figure 97 : Récapitulatif des résultats de prédiction par IRTF et UV-visible pour la modélisation globale (fermentation, élevage et vins commerciaux)

#### ❖ Perspectives générales

Les résultats de prédictions obtenues pour la concentration et la composition en anthocyanes et en tanins des différentes matrices étudiées nous ont permis de mettre en évidence les paramètres exploitables, et ceux dont la modélisation nécessite d'être approfondie. Pour augmenter la robustesse de la calibration, il est nécessaire d'ajouter de nouveaux échantillons avec une forte variabilité par rapport à ceux déjà utilisés, comme des raisins, moûts et vins issus de cépages résistants aux maladies cryptogamiques ou de différentes fermentations comme la macération carbonique par exemple. Afin de mettre à l'épreuve les modèles obtenus, il serait également intéressant de les soumettre à des vins dopés en anthocyanes et en tanins, et d'étudier également l'influence de paramètres chimiques tels que la concentration en sucres, le pH ou les concentrations en différents acides organiques du vin. Nous avons mis en évidence que la combinaison de l'IRTF et de l'UV-visible peut augmenter la robustesse de la prédiction des modèles. Il est possible de rechercher si l'ajout d'autres technologies analytiques, telles que le proche infrarouge, l'infrarouge à réflectance totale atténuée ou la fluorescence peuvent également augmenter la prédiction. Cependant, ces techniques sont complexes et demandent un appareillage coûteux, qui ne sera pas forcément aussi bien adapté au milieu vitivinicole que l'IRTF et l'UV-visible.

---

## **Références bibliographiques**

---



1. Cheynier, V., Duenas-Paton, M., Salas, E., Maury, C., Souquet, J. M., Sarni-Manchado, P., & Fulcrand, H. (2006). Structure and properties of wine pigments and tannins. *American Journal of Enology and Viticulture*, 57(3), 298-305.
2. Noble, A. C. (1998). Why do wines taste bitter and feel astringent?.
3. Vidal, S., Francis, L., Guyot, S., Marnet, N., Kwiatkowski, M., Gawel, R., ... & Waters, E. J. (2003). The mouth-feel properties of grape and apple proanthocyanidins in a wine-like medium. *Journal of the Science of Food and Agriculture*, 83(6), 564-573.
4. Mazza, G., & Francis, F. J. (1995). Anthocyanins in grapes and grape products. *Critical Reviews in Food Science & Nutrition*, 35(4), 341-371.
5. Somers, T. C. (1971). The polymeric nature of wine pigments. *Phytochemistry*, 10(9), 2175-2186.
6. Cozzolino, D., Cynkar, W. U., Shah, N., & Smith, P. (2011). Multivariate data analysis applied to spectroscopy: Potential application to juice and fruit quality. *Food Research International*, 44(7), 1888-1896.
7. Haaland, D. M., & Thomas, E. V. (1988). Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical chemistry*, 60(11), 1193-1202.
8. Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems*, 58(2), 109-130.
9. Skogerson, K., Downey, M., Mazza, M., & Boulton, R. (2007). Rapid determination of phenolic components in red wines from UV-visible spectra and the method of partial least squares. *American journal of enology and viticulture*, 58(3), 318-325.
10. Martelo-Vidal, M. J., & Vázquez, M. (2014). Determination of polyphenolic compounds of red wines by UV-VIS-NIR spectroscopy and chemometrics tools. *Food chemistry*, 158, 28-34.
11. Aleixandre-Tudo, J. L., Nieuwoudt, H., Aleixandre, J. L., & Du Toit, W. J. (2015). Robust ultraviolet-visible (UV-Vis) partial least-squares (PLS) models for Tannin quantification in red wine. *Journal of agricultural and food chemistry*, 63(4), 1088-1098.
12. Aleixandre-Tudo, J. L., Nieuwoudt, H., Olivieri, A., Aleixandre, J. L., & du Toit, W. (2018). Phenolic profiling of grapes, fermenting samples and wines using UV-Visible spectroscopy with chemometrics. *Food control*, 85, 11-22.
13. Soriano, A., Pérez-Juan, P. M., Vicario, A., González, J. M., & Pérez-Coello, M. S. (2007). Determination of anthocyanins in red wine using a newly developed method based on Fourier transform infrared spectroscopy. *Food chemistry*, 104(3), 1295-1303.
14. Cozzolino, D., Kwiatkowski, M. J., Parker, M., Cynkar, W. U., Damberg, R. G., Gishen, M., & Herderich, M. J. (2004). Prediction of phenolic compounds in red wine fermentations by visible and near infrared spectroscopy. *Analytica Chimica Acta*, 513(1), 73-80.

15. Jensen, J. S., Egebo, M., & Meyer, A. S. (2008). Identification of spectral regions for the quantification of red wine tannins with Fourier transform mid-infrared spectroscopy. *Journal of agricultural and food chemistry*, *56*(10), 3493-3499.
16. Dai, J., & Mumper, R. J. (2010). Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. *Molecules*, *15*(10), 7313-7352.
17. Beckman, C. H. (2000). Phenolic-storing cells: keys to programmed cell death and periderm formation in wilt disease resistance and in general defence responses in plants?. *Physiological and molecular plant pathology*, *57*(3), 101-110.
18. Bravo, L. (1998). Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. *Nutrition reviews*, *56*(11), 317-333.
19. Balasundram, N., Sundram, K., & Samman, S. (2006). Phenolic compounds in plants and agri-industrial by-products: Antioxidant activity, occurrence, and potential uses. *Food chemistry*, *99*(1), 191-203.
20. Dubick, M. A., & Omaye, S. T. (2001). Evidence for grape, wine and tea polyphenols as modulators of atherosclerosis and ischemic heart disease in humans. *Journal of nutraceuticals, functional & medical foods*, *3*(3), 67-93.
21. Frankel, E. N., German, J. B., Kinsella, J. E., Parks, E., & Kanner, J. (1993). Inhibition of oxidation of human low-density lipoprotein by phenolic substances in red wine. *The Lancet*, *341*(8843), 454-457.
22. Teissedre, P. L., & Landrault, N. (2000). Wine phenolics: contribution to dietary intake and bioavailability. *Food Research International*, *33*(6), 461-467.
23. Tressera-Rimbau, A., Arranz, S., Eder, M., & Vallverdú-Queralt, A. (2017). Dietary polyphenols in the prevention of stroke. *Oxidative medicine and cellular longevity*, *2017*.
24. Chen, D. I., & Dou, Q. P. (2008). Tea polyphenols and their roles in cancer prevention and chemotherapy. *International journal of molecular sciences*, *9*(7), 1196-1206.
25. He, S., Sun, C., & Pan, Y. (2008). Red wine polyphenols for cancer prevention. *International journal of molecular sciences*, *9*(5), 842-853.
26. Kim, M. J., Kim, Y. J., Park, H. J., Chung, J. H., Leem, K. H., & Kim, H. K. (2006). Apoptotic effect of red wine polyphenols on human colon cancer SNU-C4 cells. *Food and chemical toxicology*, *44*(6), 898-902.
27. Turrini, E., Ferruzzi, L., & Fimognari, C. (2015). Potential effects of pomegranate polyphenols in cancer prevention and therapy. *Oxidative medicine and cellular longevity*, *2015*.
28. Bhullar, K. S., & Rupasinghe, H. P. (2013). Polyphenols: multipotent therapeutic agents in neurodegenerative diseases. *Oxidative medicine and cellular longevity*, *2013*.
29. Ebrahimi, A., & Schluesener, H. (2012). Natural polyphenols against neurodegenerative disorders: potentials and pitfalls. *Ageing research reviews*, *11*(2), 329-345.

30. Weinreb, O., Mandel, S., Amit, T., & Youdim, M. B. (2004). Neurological mechanisms of green tea polyphenols in Alzheimer's and Parkinson's diseases. *The Journal of nutritional biochemistry*, 15(9), 506-516.
31. Fraga, C. G., Galleano, M., Verstraeten, S. V., & Oteiza, P. I. (2010). Basic biochemical mechanisms behind the health benefits of polyphenols. *Molecular aspects of medicine*, 31(6), 435-445.
32. Galleano, M., Verstraeten, S. V., Oteiza, P. I., & Fraga, C. G. (2010). Antioxidant actions of flavonoids: thermodynamic and kinetic analysis. *Archives of Biochemistry and Biophysics*, 501(1), 23-30.
33. Fraga, C. G., Oteiza, P. I., & Galleano, M. (2018). Plant bioactives and redox signaling:(-)-Epicatechin as a paradigm. *Molecular aspects of medicine*, 61, 31-40.
34. Kim, H. S., Quon, M. J., & Kim, J. A. (2014). New insights into the mechanisms of polyphenols beyond antioxidant properties; lessons from the green tea polyphenol, epigallocatechin 3-gallate. *Redox biology*, 2, 187-195.
35. Swain, T., & Bate-Smith, E. C. (1962). Flavonoid compounds. *Comparative biochemistry*, 755-809.
36. Haslam, E., & Cai, Y. (1994). Plant polyphenols (vegetable tannins): gallic acid metabolism. *Natural product reports*, 11, 41-66..
37. Quideau, S., Deffieux, D., Douat-Casassus, C., & Pouysegu, L. (2011). Plant polyphenols: chemical properties, biological activities, and synthesis. *Angewandte Chemie International Edition*, 50(3), 586-621.
38. Nagel, C. W., & Wulf, L. W. (1979). Changes in the anthocyanins, flavonoids and hydroxycinnamic acid esters during fermentation and aging of Merlot and Cabernet Sauvignon. *American Journal of Enology and Viticulture*, 30(2), 111-116.
39. Singleton, V. L., & Draper, D. E. (1964). The transfer of polyphenolic compounds from grape seeds into wines. *American Journal of Enology and Viticulture*, 15(1), 34-40.
40. Brouillard, R., George, F., & Fougèrouse, A. (1997). Polyphenols produced during red wine ageing. *Biofactors*, 6(4), 403-410.
41. De Simón, B. F., Sanz, M., Cadahía, E., Martínez, J., Esteruelas, E., & Muñoz, A. M. (2014). Polyphenolic compounds as chemical markers of wine ageing in contact with cherry, chestnut, false acacia, ash and oak wood. *Food chemistry*, 143, 66-76.
42. Crozier, A., Jaganath, I. B., & Clifford, M. N. (2009). Dietary phenolics: chemistry, bioavailability and effects on health. *Natural product reports*, 26(8), 1001-1043.
43. Marković, D., Petranović, N. A., & Baranac, J. M. (2000). A spectrophotometric study of the copigmentation of malvin with caffeic and ferulic acids. *Journal of agricultural and food chemistry*, 48(11), 5530-5536.

44. Baur, J. A., Pearson, K. J., Price, N. L., Jamieson, H. A., Lerin, C., Kalra, A., ... & Sinclair, D. A. (2006). Resveratrol improves health and survival of mice on a high-calorie diet. *Nature*, *444*(7117), 337-342.
45. Fremont, L. (2000). Biological effects of resveratrol. *Life sciences*, *66*(8), 663-673.
46. Vanhoenacker, G., De Villiers, A., Lazou, K., De Keukeleire, D., & Sandra, P. (2001). Comparison of high-performance liquid chromatography—Mass spectroscopy and capillary electrophoresis—Mass spectroscopy for the analysis of phenolic compounds in diethyl ether extracts of red wines. *Chromatographia*, *54*(5-6), 309-315.
47. Chira, K., Suh, J. H., Saucier, C., & Teissèdre, P. L. (2008). Les polyphénols du raisin. *Phytothérapie*, *6*(2), 75-82.
48. Moreno-Arribas, M. V., & Polo, M. C. (2009). Volatile and Aroma Compounds. In *Wine Chemistry and Biochemistry* (pp. 249-249). Springer, New York, NY.
49. Galvin, C. (1993). *Etude de certaines réactions de dégradation des anthocyanes et de leur condensation avec les flavonols: conséquences sur la couleur des vins* (Doctoral dissertation, Bordeaux 2).
50. Kennedy, J. A., Saucier, C., & Glories, Y. (2006). Grape and wine phenolics: History and perspective. *American Journal of Enology and Viticulture*, *57*(3), 239-248.
51. Ong, B. Y., & Nagel, C. W. (1978). High-pressure liquid chromatographic analysis of hydroxycinnamic acid—tartaric acid esters and their glucose esters in *Vitis vinifera*. *Journal of Chromatography A*, *157*, 345-355.
52. Monagas, M., Bartolomé, B., & Gómez-Cordovés, C. (2005). Updated knowledge about the presence of phenolic compounds in wine. *Critical reviews in food science and nutrition*, *45*(2), 85-118.
53. Somers, T. C., Vérette, E., & Pocock, K. F. (1987). Hydroxycinnamate esters of *Vitis vinifera*: changes during white vinification, and effects of exogenous enzymic hydrolysis. *Journal of the Science of Food and Agriculture*, *40*(1), 67-78.
54. Darias-Martín, J., Martín-Luis, B., Carrillo-López, M., Lamuela-Raventós, R., Díaz-Romero, C., & Boulton, R. (2002). Effect of caffeic acid on the color of red wine. *Journal of agricultural and food chemistry*, *50*(7), 2062-2067.
55. Rentzsch, M., Schwarz, M., & Winterhalter, P. (2007). Pyranoanthocyanins—an overview on structures, occurrence, and pathways of formation. *Trends in Food Science & Technology*, *18*(10), 526-534.
56. Schwarz, M., Wabnitz, T. C., & Winterhalter, P. (2003). Pathway leading to the formation of anthocyanin–vinylphenol adducts and related pigments in red wines. *Journal of Agricultural and Food Chemistry*, *51*(12), 3682-3687.
57. Chatonnet, P. (1995). *Influence des procédés de tonnellerie et des conditions d'élevage sur la composition et la qualité des vins élevés en fûts de chêne* (Doctoral dissertation, Bordeaux 2).

58. Hixson, J. L., Sleep, N. R., Capone, D. L., Elsey, G. M., Curtin, C. D., Sefton, M. A., & Taylor, D. K. (2012). Hydroxycinnamic acid ethyl esters as precursors to ethylphenols in wine. *Journal of agricultural and food chemistry*, *60*(9), 2293-2298.
59. Mattivi, F., Reniero, F., & Korhammer, S. (1995). Isolation, characterization, and evolution in red wine vinification of resveratrol monomers. *Journal of Agricultural and Food Chemistry*, *43*(7), 1820-1823.
60. Chong, J., Poutaraud, A., & Hugueney, P. (2009). Metabolism and roles of stilbenes in plants. *Plant science*, *177*(3), 143-155.
61. Sun, B., Ribes, A. M., Leandro, M. C., Belchior, A. P., & Spranger, M. I. (2006). Stilbenes: quantitative extraction from grape skins, contribution of grape solids to wine and variation during wine maturation. *Analytica chimica acta*, *563*(1-2), 382-390.
62. Flamini, R., Mattivi, F., Rosso, M. D., Arapitsas, P., & Bavaresco, L. (2013). Advanced knowledge of three important classes of grape phenolics: anthocyanins, stilbenes and flavonols. *International journal of molecular sciences*, *14*(10), 19651-19669.
63. Jean-Denis, J. B., Pezet, R., & Tabacchi, R. (2006). Rapid analysis of stilbenes and derivatives from downy mildew-infected grapevine leaves by liquid chromatography–atmospheric pressure photoionisation mass spectrometry. *Journal of Chromatography A*, *1112*(1-2), 263-268.
64. Goldberg, D. M., Ng, E., Yan, J., Karumanchiri, A., Soleas, G. J., & Diamandis, E. P. (1996). Regional differences in resveratrol isomer concentrations of wines from various cultivars. *Journal of Wine Research*, *7*(1), 13-24.
65. Jeandet, P., Bessis, R., & Gautheron, B. (1991). The production of resveratrol (3, 5, 4'-trihydroxystilbene) by grape berries in different developmental stages. *American Journal of Enology and Viticulture*, *42*(1), 41-46.
66. Roldán, A., Palacios, V., Caro, I., & Pérez, L. (2003). Resveratrol content of Palomino fino grapes: influence of vintage and fungal infection. *Journal of agricultural and food chemistry*, *51*(5), 1464-1468.
67. Jeandet, P., Bessis, R., Sbaghi, M., & Meunier, P. (1995). Production of the phytoalexin resveratrol by grapes as a response to Botrytis attack under natural conditions. *Journal of Phytopathology*, *143*(3), 135-139.
68. Guerrero, R. F., Valls-Fonayet, J., Richard, T., & Cantos-Villar, E. (2020). A rapid quantification of stilbene content in wine by ultra-high pressure liquid chromatography–Mass spectrometry. *Food Control*, *108*, 106821.
69. Moreno-Labanda, J. F., Mallavia, R., Pérez-Fons, L., Lizama, V., Saura, D., & Micol, V. (2004). Determination of piceid and resveratrol in Spanish wines deriving from Monastrell (*Vitis vinifera* L.) grape variety. *Journal of agricultural and food chemistry*, *52*(17), 5396-5403.

70. Montsko, G., Nikfardjam, M. S. P., Szabo, Z., Boddi, K., Lorand, T., Ohmacht, R., & Mark, L. (2008). Determination of products derived from trans-resveratrol UV photoisomerisation by means of HPLC–APCI-MS. *Journal of Photochemistry and Photobiology A: Chemistry*, 196(1), 44-50.
71. Pastor, R. F., Restani, P., Di Lorenzo, C., Orgiu, F., Teissedre, P. L., Stockley, C., ... & Iermoli, R. H. (2019). Resveratrol, human health and winemaking perspectives. *Critical reviews in food science and nutrition*, 59(8), 1237-1255.
72. He, S., & Yan, X. (2013). From resveratrol to its derivatives: new sources of natural antioxidant. *Current medicinal chemistry*, 20(8), 1005-1017.
73. Venturini, C. D., Merlo, S., Souto, A. A., Fernandes, M. D. C., Gomez, R., & Rhoden, C. R. (2010). Resveratrol and red wine function as antioxidants in the nervous system without cellular proliferative effects during experimental diabetes. *Oxidative Medicine and Cellular Longevity*, 3(6), 434-441.
74. Baur, J. A., & Sinclair, D. A. (2006). Therapeutic potential of resveratrol: the in vivo evidence. *Nature reviews Drug discovery*, 5(6), 493-506.
75. Penumathsa, S. V., & Maulik, N. (2009). Resveratrol: a promising agent in promoting cardioprotection against coronary heart disease. *Canadian journal of physiology and pharmacology*, 87(4), 275-286.
76. Aggarwal, B. B., Bhardwaj, A., Aggarwal, R. S., Seeram, N. P., Shishodia, S., & Takada, Y. (2004). Role of resveratrol in prevention and therapy of cancer: preclinical and clinical studies. *Anticancer research*, 24(5A), 2783-2840.
77. Jang, M., Cai, L., Udeani, G. O., Slowing, K. V., Thomas, C. F., Beecher, C. W., ... & Pezzuto, J. M. (1997). Cancer chemopreventive activity of resveratrol, a natural product derived from grapes. *Science*, 275(5297), 218-220.
78. Alarcon De La Lastra, C., & Villegas, I. (2005). Resveratrol as an anti-inflammatory and anti-aging agent: Mechanisms and clinical implications. *Molecular nutrition & food research*, 49(5), 405-430.
79. Rahman, I., Biswas, S. K., & Kirkham, P. A. (2006). Regulation of inflammation and redox signaling by dietary polyphenols. *Biochemical pharmacology*, 72(11), 1439-1452.
80. Lin, C. T., Sun, X. Y., & Lin, A. X. (2016). Supplementation with high-dose trans-resveratrol improves ultrafiltration in peritoneal dialysis patients: a prospective, randomized, double-blind study. *Renal failure*, 38(2), 214-221.
81. Wong, R. H., Raederstorff, D., & Howe, P. R. (2016). Acute resveratrol consumption improves neurovascular coupling capacity in adults with type 2 diabetes mellitus. *Nutrients*, 8(7), 425.
82. Carando, S., Teissedre, P. L., Pascual-Martinez, L., & Cabanis, J. C. (1999). Levels of flavan-3-ols in French wines. *Journal of agricultural and food chemistry*, 47(10), 4161-4166.

83. Fang, F., Li, J. M., Zhang, P., Tang, K., Wang, W., Pan, Q. H., & Huang, W. D. (2008). Effects of grape variety, harvest date, fermentation vessel and wine ageing on flavonoid concentration in red wines. *Food Research International*, 41(1), 53-60.
84. Garrido, J., & Borges, F. (2013). Wine and grape polyphenols—A chemical perspective. *Food research international*, 54(2), 1844-1858.
85. Tsao, R. (2010). Chemistry and biochemistry of dietary polyphenols. *Nutrients*, 2(12), 1231-1246.
86. Jandera, P., Škeříková, V., Řehová, L., Hájek, T., Baldriánová, L., Škopová, G., ... & Horna, A. (2005). RP-HPLC analysis of phenolic compounds and flavonoids in beverages and plant extracts using a CoulArray detector. *Journal of separation science*, 28(9-10), 1005-1022.
87. Pereira, G. E., Gaudillere, J. P., Pieri, P., Hilbert, G., Maucourt, M., Deborde, C., ... & Rolin, D. (2006). Microclimate influence on mineral and metabolic profiles of grape berries. *Journal of Agricultural and Food Chemistry*, 54(18), 6765-6775.
88. Montealegre, R. R., Peces, R. R., Vozmediano, J. C., Gascueña, J. M., & Romero, E. G. (2006). Phenolic compounds in skins and seeds of ten grape *Vitis vinifera* varieties grown in a warm climate. *Journal of Food Composition and Analysis*, 19(6-7), 687-693.
89. Makris, D. P., Kallithraka, S., & Kefalas, P. (2006). Flavonols in grapes, grape products and wines: Burden, profile and influential parameters. *Journal of Food Composition and Analysis*, 19(5), 396-404.
90. Mattivi, F., Guzzon, R., Vrhovsek, U., Stefanini, M., & Velasco, R. (2006). Metabolite profiling of grape: flavonols and anthocyanins. *Journal of agricultural and food chemistry*, 54(20), 7692-7702.
91. Jeffery, D. W., Parker, M., & Smith, P. A. (2008). Flavonol composition of Australian red and white wines determined by high-performance liquid chromatography. *Australian Journal of Grape and Wine Research*, 14(3), 153-161.
92. Lukić, I., Radeka, S., Budić-Leto, I., Bubola, M., & Vrhovsek, U. (2019). Targeted UPLC-QqQ-MS/MS profiling of phenolic compounds for differentiation of monovarietal wines and corroboration of particular varietal typicity concepts. *Food chemistry*, 300, 125251.
93. Zoecklein, B., Fugelsang, K. C., Gump, B. H., & Nury, F. S. (2013). *Wine analysis and production*. Springer Science & Business Media.
94. Singleton, V. L. (1966). The total phenolic content of grape berries during the maturation of several varieties. *American Journal of Enology and Viticulture*, 17(2), 126-134.
95. Lorrain, B., Chira, K., & Teissedre, P. L. (2011). Phenolic composition of Merlot and Cabernet-Sauvignon grapes from Bordeaux vineyard for the 2009-vintage: Comparison to 2006, 2007 and 2008 vintages. *Food Chemistry*, 126(4), 1991-1999.
96. de Andrés-de Prado, R., Yuste-Rojas, M., Sort, X., Andrés-Lacueva, C., Torres, M., & Lamuela-Raventós, R. M. (2007). Effect of soil type on wines produced from *Vitis vinifera* L.

- cv. Grenache in commercial vineyards. *Journal of agricultural and food chemistry*, 55(3), 779-786.
97. Downey, M. O., Harvey, J. S., & Robinson, S. P. (2003). Analysis of tannins in seeds and skins of Shiraz grapes throughout berry development. *Australian Journal of Grape and Wine Research*, 9(1), 15-27.
  98. Czochanska, Z., Foo, L. Y., & Porter, L. J. (1979). Compositional changes in lower molecular weight flavans during grape maturation. *Phytochemistry*, 18(11), 1819-1822.
  99. Jaworski, A. W., & Lee, C. Y. (1987). Fractionation and HPLC determination of grape phenolics. *Journal of Agricultural and Food Chemistry*, 35(2), 257-259.
  100. Su, C. T., & Singleton, V. L. (1969). Identification of three flavan-3-ols from grapes. *Phytochemistry*, 8(8), 1553-1558.
  101. Lee, C. Y., & Jaworski, A. W. (1990). Identification of some phenolics in white grapes. *American journal of enology and viticulture*, 41(1), 87-89.
  102. Ma, W., Waffo-Téguo, P., Jourdes, M., Li, H., & Teissedre, P. L. (2018). First evidence of epicatechin vanillate in grape seed and red wine. *Food chemistry*, 259, 304-310.
  103. Cheynier, V., Prieur, C., Guyot, S., Rigaud, J., & Moutounet, M. (1997). The structures of tannins in grapes and wines and their interactions with proteins.
  104. Gagné, S., Saucier, C., & Gény, L. (2006). Composition and cellular localization of tannins in Cabernet Sauvignon skins during growth. *Journal of agricultural and food chemistry*, 54(25), 9465-9471.
  105. Amrani Joutei, K., Glories, Y., & Mercier, M. (1994). Localisation des tanins dans la pellicule de baie de raisin. *Vitis*, 33(3), 133-138.
  106. Geny, L., Saucier, C., Bracco, S., Daviaud, F., & Glories, Y. (2003). Composition and cellular localization of tannins in grape seeds during maturation. *Journal of Agricultural and Food Chemistry*, 51(27), 8051-8054.
  107. Thorngate, J. H., & Singleton, V. L. (1994). Localization of procyanidins in grape seeds. *American Journal of Enology and Viticulture*, 45(2), 259-262.
  108. Aron, P. M., & Kennedy, J. A. (2008). Flavan-3-ols: Nature, occurrence and biological activity. *Molecular nutrition & food research*, 52(1), 79-104.
  109. Peyrot des Gachons, C., & Kennedy, J. A. (2003). Direct method for determining seed and skin proanthocyanidin extraction into red wine. *Journal of agricultural and food chemistry*, 51(20), 5877-5881.
  110. Santos-Buelga, C., Francia-Aricha, E. M., & Escribano-Bailón, M. T. (1995). Comparative flavan-3-ol composition of seeds from different grape varieties. *Food Chemistry*, 53(2), 197-201.

111. Da Silva, J. M. R., Rigaud, J., Cheynier, V., Cheminat, A., & Moutounet, M. (1991). Procyanidin dimers and trimers from grape seeds. *Phytochemistry*, *30*(4), 1259-1264.
112. Batesmith, E., & Swain, T. (1965, January). Recent developments in chemotaxonomy of flavonoid compounds. In *Lloydia* (Vol. 28, No. 4).
113. Riou, V., Vernhet, A., Doco, T., & Moutounet, M. (2002). Aggregation of grape seed tannins in model wine—effect of wine polysaccharides. *Food hydrocolloids*, *16*(1), 17-23.
114. Ribéreau-Gayon, P., & Stonestreet, E. (1966). Dosage des tanins du vin rouge et détermination de leur structure. *Chimie analytique*, *48*(4).
115. He, F., Pan, Q. H., Shi, Y., & Duan, C. Q. (2008). Chemical synthesis of proanthocyanidins in vitro and their reactions in aging wines. *Molecules*, *13*(12), 3007-3032.
116. Appeldoorn, M. M., Sanders, M., Vincken, J. P., Cheynier, V., Le Guernevé, C., Hollman, P. C., & Gruppen, H. (2009). Efficient isolation of major procyanidin A-type dimers from peanut skins and B-type dimers from grape seeds. *Food chemistry*, *117*(4), 713-720.
117. Jacques, D., Haslam, E., Bedford, G. R., & Greatbanks, D. (1973). Structure of the dimeric proanthocyanidin-A2 and its derivatives. *Journal of the Chemical Society, Chemical Communications*, (15), 518-520.
118. Kyraleou, M., Kallithraka, S., Theodorou, N., Teissedre, P. L., Kotseridis, Y., & Koundouras, S. (2017). Changes in tannin composition of Syrah grape skins and seeds during fruit ripening under contrasting water conditions. *Molecules*, *22*(9), 1453.
119. Rousserie, P., Rabot, A., & Geny-Denis, L. (2019). From flavanols biosynthesis to wine tannins: What place for grape seeds?. *Journal of agricultural and food chemistry*, *67*(5), 1325-1343.
120. Delcambre, A., & Saucier, C. (2013). High-throughput oenomics: shotgun polyphenomics of wines. *Analytical chemistry*, *85*(20), 9736-9741.
121. Lisjak, K., Lelova, Z., Žigon, U., Bolta, Š. V., Teissedre, P. L., & Vanzo, A. (2020). Effect of extraction time on content, composition and sensory perception of proanthocyanidins in wine-like medium and during industrial fermentation of Cabernet Sauvignon. *Journal of the Science of Food and Agriculture*, *100*(5), 1887-1896.
122. Zeng, L., Pons-Mercadé, P., Richard, T., Krisa, S., Teissèdre, P. L., & Jourdes, M. (2019). Crown procyanidin tetramer: A procyanidin with an unusual cyclic skeleton with a potent protective effect against amyloid- $\beta$ -induced toxicity. *Molecules*, *24*(10), 1915.
123. Merkytė, V., Longo, E., Jourdes, M., Jouin, A., Teissedre, P. L., & Boselli, E. (2020). High-Performance Liquid Chromatography–Hydrogen/Deuterium Exchange–High-Resolution Mass Spectrometry Partial Identification of a Series of Tetra- and Pentameric Cyclic Procyanidins and Prodelphinidins in Wine Extracts. *Journal of agricultural and food chemistry*, *68*(11), 3312-3321.

124. Zeng, L. (2015). *Étude de la composition macromoléculaire du raisin et des vins: impact sur la qualité sensorielle* (Doctoral dissertation, Bordeaux).
125. Longo, E., Merkyte, V., Rossetti, F., Teissedre, P. L., Jourdes, M., & Boselli, E. (2018). Relative abundances of novel cyclic prodelphinidins in wine depending on the grape variety. *Journal of Mass Spectrometry*, *53*(11), 1116-1125.
126. Longo, E., Rossetti, F., Jouin, A., Teissedre, P. L., Jourdes, M., & Boselli, E. (2019). Distribution of crown hexameric procyanidin and its tetrameric and pentameric congeners in red and white wines. *Food chemistry*, *299*, 125125.
127. Ikan, R. (1991). *Natural products: a laboratory guide*. Academic Press.
128. Heredia, F. J., Francia-Aricha, E. M., Rivas-Gonzalo, J. C., Vicario, I. M., & Santos-Buelga, C. (1998). Chromatic characterization of anthocyanins from red grapes—I. pH effect. *Food Chemistry*, *63*(4), 491-498.
129. Burns, J., Mullen, W., Landrault, N., Teissedre, P. L., Lean, M. E., & Crozier, A. (2002). Variations in the profile and content of anthocyanins in wines made from Cabernet Sauvignon and hybrid grapes. *Journal of agricultural and food chemistry*, *50*(14), 4096-4102.
130. Mulinacci, N., Santamaria, A. R., Giaccherini, C., Innocenti, M., Valletta, A., Ciolfi, G., & Pasqua, G. (2008). Anthocyanins and flavan-3-ols from grapes and wines of *Vitis vinifera* cv. Cesanese d’Affile. *Natural product research*, *22*(12), 1033-1039.
131. Nikfardjam, M. S. P., Márk, L., Avar, P., Figler, M., & Ohmacht, R. (2006). Polyphenols, anthocyanins, and trans-resveratrol in red wines from the Hungarian Villány region. *Food Chemistry*, *98*(3), 453-462.
132. Pecket, R. C., & Small, C. J. (1980). Occurrence, location and development of anthocyanoplasts. *Phytochemistry*, *19*(12), 2571-2576.
133. Guan, L., Li, J. H., Fan, P. G., Li, S. H., Fang, J. B., Dai, Z. W., ... & Wu, B. H. (2014). Regulation of anthocyanin biosynthesis in tissues of a teinturier grape cultivar under sunlight exclusion. *American Journal of Enology and Viticulture*, *65*(3), 363-374.
134. He, F., Liang, N. N., Duan, C. Q., Wang, J., & Reeves, M. J. (2013, April). Accumulation of anthocyanins in the skin and pulp of five different teinturier grapes. In *IX International Symposium on Grapevine Physiology and Biotechnology 1157* (pp. 73-82).
135. Mazza, G., & Miniati, E. (1993). Grapes. *Anthocyanins in fruits, vegetables and grains*, 149-199.
136. Alcalde-Eon, C., Escribano-Bailón, M. T., Santos-Buelga, C., & Rivas-Gonzalo, J. C. (2006). Changes in the detailed pigment composition of red wine during maturity and ageing: A comprehensive study. *Analytica Chimica Acta*, *563*(1-2), 238-254.
137. Castillo-Munoz, N., Fernandez-Gonzalez, M., Gomez-Alonso, S., Garcia-Romero, E., & Hermosin-Gutierrez, I. (2009). Red-color related phenolic composition of Garnacha Tintorera

- (*Vitis vinifera* L.) grapes and red wines. *Journal of Agricultural and Food Chemistry*, 57(17), 7883-7891.
138. He, F., Liang, N. N., Mu, L., Pan, Q. H., Wang, J., Reeves, M. J., & Duan, C. Q. (2012). Anthocyanins and their variation in red wines I. Monomeric anthocyanins and their color expression. *Molecules*, 17(2), 1571-1601.
  139. HE, F., He, J. J., Pan, Q. H., & Duan, C. Q. (2010). Mass-spectrometry evidence confirming the presence of pelargonidin-3-O-glucoside in the berry skins of Cabernet Sauvignon and Pinot Noir (*Vitis vinifera* L.). *Australian Journal of Grape and Wine Research*, 16(3), 464-468.
  140. Yıldırım, S., Kadioğlu, A., Sağlam, A., Yaşar, A., & Sellitepe, H. E. (2016). Fast determination of anthocyanins and free pelargonidin in fruits, fruit juices, and fruit wines by high-performance liquid chromatography using a core-shell column. *Journal of separation science*, 39(20), 3927-3935.
  141. Gonzalez-San Jose, M. L., Santa-Maria, G., & Diez, C. (1990). Anthocyanins as parameters for differentiating wines by grape variety, wine-growing region, and wine-making methods. *Journal of food composition and analysis*, 3(1), 54-66.
  142. Vivas De Gaulejac, N., Nonier, M. F., Guerra, C., & Vivas, N. (2001). Anthocyanin in grape skins during maturation of *Vitis vinifera* L. cv. Cabernet sauvignon and Merlot noir from different Bordeaux terroirs. *Journal International des Sciences de la Vigne et du Vin*, 35(3), 149-156.
  143. Ehrhardt, C., Arapitsas, P., Stefanini, M., Flick, G., & Mattivi, F. (2014). Analysis of the phenolic composition of fungus-resistant grape varieties cultivated in Italy and Germany using UHPLC-MS/MS. *Journal of mass spectrometry*, 49(9), 860-869.
  144. Dimitrovska, M., Bocevska, M., Dimitrovski, D., & Murkovic, M. (2011). Anthocyanin composition of Vranec, Cabernet Sauvignon, Merlot and Pinot Noir grapes as indicator of their varietal differentiation. *European food research and technology*, 232(4), 591-600.
  145. Brouillard, R., Delaporte, B., & Dubois, J. E. (1978). Chemistry of anthocyanin pigments. 3. Relaxation amplitudes in pH-jump experiments. *Journal of the American Chemical Society*, 100(19), 6202-6205.
  146. Glories, Y. (1984). La couleur des vins rouges. Ire partie: les équilibres des anthocyanes et des tanins. *OENO One*, 18(3), 195-217.
  147. Andorrà, I., Landi, S., Mas, A., Guillamón, J. M., & Esteve-Zarzoso, B. (2008). Effect of oenological practices on microbial populations using culture-independent techniques. *Food Microbiology*, 25(7), 849-856.
  148. Danilewicz, J. C., Secombe, J. T., & Whelan, J. (2008). Mechanism of interaction of polyphenols, oxygen, and sulfur dioxide in model wine and wine. *American journal of enology and viticulture*, 59(2), 128-136.
  149. Barbe, J. C. (2000). *La combinaison du dioxyde de soufre dans les moûts et vins issus de raisins botrytisés: rôle des bactéries acétiques* (Doctoral dissertation, Bordeaux 2).

150. Danilewicz, J. C. (2007). Interaction of sulfur dioxide, polyphenols, and oxygen in a wine-model system: Central role of iron and copper. *American Journal of Enology and Viticulture*, 58(1), 53-60.
151. Berké, B., Chèze, C., Vercauteren, J., & Deffieux, G. (1998). Bisulfite addition to anthocyanins: revisited structures of colourless adducts. *Tetrahedron Letters*, 39(32), 5771-5774.
152. Quideau, S., Jourdes, M., Lefeuvre, D., Pardon, P., Saucier, C., Teissedre, P. L., & Glories, Y. (2010). Ellagitannins-an underestimated class of plant polyphenols: chemical reactivity of C-glucosidic ellagitannins in relation to wine chemistry and biological activity. *Recent advances in polyphenol research, Volume 2*, 81-137.
153. Quideau, S., & Feldman, K. S. (1996). Ellagitannin chemistry. *Chemical Reviews*, 96(1), 475-504.
154. Jourdes, M., Pouységu, L., Deffieux, D., Teissedre, P. L., & Quideau, S. (2013). Hydrolyzable tannins: gallotannins and ellagitannins. *Natural Products*, 1975-2010.
155. Haddock, E. A., Gupta, R. K., Al-Shafi, S. M., Haslam, E., & Magnolato, D. (1982). The metabolism of gallic acid and hexahydroxydiphenic acid in plants. Part 1. Introduction. Naturally occurring galloyl esters. *Journal of the Chemical Society, Perkin Transactions 1*, 2515-2524.
156. Niemetz, R., & Gross, G. G. (2005). Enzymology of gallotannin and ellagitannin biosynthesis. *Phytochemistry*, 66(17), 2001-2011.
157. Gombau, J., Vignault, A., Pascual, O., Gomez-Alonso, S., Gracia-Romero, E., Hermosin, I., ... & Zamora, F. (2019). Influence of oenological tannins on malvidin-3-O-monoglucoside copigmentation in a model wine solution. *Oeno One*, 53(3).
158. Vignault, A., Gombau, J., Pascual, O., Jourdes, M., Moine, V., Canals, J. M., ... & Teissedre, P. L. (2019). Copigmentation of malvidin-3-O-monoglucoside by oenological tannins: Incidence on wine model color in function of botanical origin, pH and ethanol content. *Molecules*, 24(8), 1448.
159. Magalhães, L. M., Ramos, I. I., Reis, S., & Segundo, M. A. (2014). Antioxidant profile of commercial oenological tannins determined by multiple chemical assays. *Australian journal of grape and wine research*, 20(1), 72-79.
160. Vignault, A., Gombau, J., Jourdes, M., Moine, V., Canals, J. M., Fermaud, M., ... & Teissedre, P. L. (2020). Oenological tannins to prevent Botrytis cinerea damage in grapes and musts: Kinetics and electrophoresis characterization of laccase. *Food chemistry*, 316, 126334.
161. Vignault, A., Pascual, O., Jourdes, M., Moine, V., Fermaud, M., Roudet, J., ... & Zamora, F. (2019). Impact of enological tannins on laccase activity. *Oeno One*, 53(1), 27-38.
162. González-Centeno, M. R., Chira, K., & Teissedre, P. L. (2016). Ellagitannin content, volatile composition and sensory profile of wines from different countries matured in oak barrels subjected to different toasting methods. *Food chemistry*, 210, 500-511.

163. Jourdes, M., Michel, J., Saucier, C., Quideau, S., & Teissedre, P. L. (2011). Identification, amounts, and kinetics of extraction of C-glucosidic ellagitannins during wine aging in oak barrels or in stainless steel tanks with oak chips. *Analytical and bioanalytical chemistry*, 401(5), 1531-1539.
164. Maga, J. A. (1989). The contribution of wood to the flavor of alcoholic beverages. *Food Reviews International*, 5(1), 39-99.
165. Singleton, V. L. (1995). Maturation of wines and spirits: comparisons, facts, and hypotheses. *American Journal of Enology and Viticulture*, 46(1), 98-115.
166. Vivas, N., & Glories, Y. (1996). Role of oak wood ellagitannins in the oxidation process of red wines during aging. *American journal of enology and viticulture*, 47(1), 103-107.
167. Mosedale, J. R., Puech, J. L., & Feuillat, F. (1999). The influence on wine flavor of the oak species and natural variation of heartwood components. *American journal of enology and viticulture*, 50(4), 503-512.
168. De Simón, B. F., Hernández, T., Cadahía, E., Dueñas, M., & Estrella, I. (2003). Phenolic compounds in a Spanish red wine aged in barrels made of Spanish, French and American oak wood. *European Food Research and Technology*, 216(2), 150-156.
169. García-Estévez, I., Escribano-Bailón, M. T., Rivas-Gonzalo, J. C., & Alcalde-Eon, C. (2017). Effect of the type of oak barrels employed during ageing on the ellagitannin profile of wines. *Australian Journal of Grape and Wine Research*, 23(3), 334-341.
170. Doussot, F., De Jéso, B., Quideau, S., & Pardon, P. (2002). Extractives content in cooperage oak wood during natural seasoning and toasting; influence of tree species, geographic location, and single-tree effects. *Journal of Agricultural and food chemistry*, 50(21), 5955-5961.
171. Masson, G., Moutounet, M., & Puech, J. L. (1995). Ellagitannin content of oak wood as a function of species and of sampling position in the tree. *American journal of enology and viticulture*, 46(2), 262-268.
172. Matricardi, L., & Waterhouse, A. L. (1999). Influence of toasting technique on color and ellagitannins of oak wood in barrel making. *American Journal of Enology and Viticulture*, 50(4), 519-526.
173. Michel, J., Jourdes, M., Silva, M. A., Giordanengo, T., Mourey, N., & Teissedre, P. L. (2011). Impact of concentration of ellagitannins in oak wood on their levels and organoleptic influence in red wine. *Journal of Agricultural and Food Chemistry*, 59(10), 5677-5683.
174. Rasines-Perea, Z., Jacquet, R., Jourdes, M., Quideau, S., & Teissedre, P. L. (2019). Ellagitannins and flavano-ellagitannins: Red wines tendency in different areas, barrel origin and ageing time in barrel and bottle. *Biomolecules*, 9(8), 316.
175. Saucier, C., Jourdes, M., Glories, Y., & Quideau, S. (2006). Extraction, detection, and quantification of flavano-ellagitannins and ethylvescalagin in a Bordeaux red wine aged in oak barrels. *Journal of Agricultural and Food Chemistry*, 54(19), 7349-7354.

176. Chassaing, S., Lefeuvre, D., Jacquet, R., Jourdes, M., Ducasse, L., Galland, S., ... & Quideau, S. (2010). Physicochemical studies of new anthocyano-ellagitannin hybrid pigments: about the origin of the influence of oak C-glycosidic ellagitannins on wine color.
177. Edwin, H. (1980). In vino veritas: oligomeric procyanidins and the ageing of red wines. *Phytochemistry*, *19*(12), 2577-2582.
178. Matthews, S., Mila, I., Scalbert, A., Pollet, B., Lapiere, C., Hervé du Penhoat, C. L., ... & Donnelly, D. M. (1997). Method for estimation of proanthocyanidins based on their acid depolymerization in the presence of nucleophiles. *Journal of Agricultural and Food Chemistry*, *45*(4), 1195-1201.
179. Vidal, S., Cartalade, D., Souquet, J. M., Fulcrand, H., & Cheynier, V. (2002). Changes in proanthocyanidin chain length in winelike model solutions. *Journal of Agricultural and Food Chemistry*, *50*(8), 2261-2266.
180. Ribéreau-Gayon, P. (1982). The anthocyanins of grapes and wines. *Anthocyanins as food colors*, *6*, 214-215.
181. Guyot, S., Vercauteren, J., & Cheynier, V. (1996). Structural determination of colourless and yellow dimers resulting from (+)-catechin coupling catalysed by grape polyphenoloxidase. *Phytochemistry*, *42*(5), 1279-1288.
182. Kondo, K., Kurihara, M., Fukuhara, K., Tanaka, T., Suzuki, T., Miyata, N., & Toyoda, M. (2000). Conversion of procyanidin B-type (catechin dimer) to A-type: evidence for abstraction of C-2 hydrogen in catechin during radical oxidation. *Tetrahedron Letters*, *41*(4), 485-488.
183. Bishop, P. D., & Nagel, C. W. (1984). Characterization of the condensation product of malvidin 3, 5-diglucoside and catechin. *Journal of agricultural and food chemistry*, *32*(5), 1022-1026.
184. Jurd, L. (1967). Anthocyanidins and related compounds—XI: Catechin-flavylium salt condensation reactions. *Tetrahedron*, *23*(3), 1057-1064.
185. Es-Safi, N. E., Fulcrand, H., Cheynier, V., & Moutounet, M. (1999). Studies on the acetaldehyde-induced condensation of (–)-epicatechin and malvidin 3-O-glucoside in a model solution system. *Journal of Agricultural and Food Chemistry*, *47*(5), 2096-2102.
186. Hayasaka, Y., & Kennedy, J. A. (2003). Mass spectrometric evidence for the formation of pigmented polymers in red wine. *Australian Journal of Grape and Wine Research*, *9*(3), 210-220.
187. He, F., Liang, N. N., Mu, L., Pan, Q. H., Wang, J., Reeves, M. J., & Duan, C. Q. (2012). Anthocyanins and their variation in red wines II. Anthocyanin derived pigments and their color evolution. *Molecules*, *17*(2), 1483-1519.
188. Santos-Buelga, C., Bravo-Haro, S., & Rivas-Gonzalo, J. C. (1995). Interactions between catechin and malvidin-3-monoglucoside in model solutions. *Zeitschrift für Lebensmittel-Untersuchung und Forschung*, *201*(3), 269-274.

189. Zeng, L., Teissèdre, P. L., & Jourdes, M. (2016). Structures of polymeric pigments in red wine and their derived quantification markers revealed by high-resolution quadrupole time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 30(1), 81-88.
190. Han, G., Webb, M. R., & Waterhouse, A. L. (2019). Acetaldehyde reactions during wine bottle storage. *Food chemistry*, 290, 208-215.
191. Romano, P., Suzzi, G., Turbanti, L., & Polsinelli, M. (1994). Acetaldehyde production in *Saccharomyces cerevisiae* wine yeasts. *FEMS Microbiology Letters*, 118(3), 213-218.
192. Wildenradt, H. L., & Singleton, V. L. (1974). The production of aldehydes as a result of oxidation of polyphenolic compounds and its relation to wine aging. *American Journal of Enology and Viticulture*, 25(2), 119-126.
193. Doco, T., Es-Safi, N. E., Cheynier, V., & Moutounet, M. (1996). Study of the acetaldehyde induced polymerisation of flavan-3-ols by liquid chromatography-ion spray mass spectrometry. *Journal of Chromatography A*, 752(1-2), 85-91.
194. Timberlake, C. F., & Bridle, P. (1976). Interactions between anthocyanins, phenolic compounds, and acetaldehyde and their significance in red wines. *American Journal of Enology and Viticulture*, 27(3), 97-105.
195. Es-Safi, N. E., Fulcrand, H., Cheynier, V., & Moutounet, M. (1999). Competition between (+)-catechin and (–)-epicatechin in acetaldehyde-induced polymerization of flavanols. *Journal of Agricultural and Food Chemistry*, 47(5), 2088-2095.
196. Atanasova, V., Fulcrand, H., Cheynier, V., & Moutounet, M. (2002). Effect of oxygenation on polyphenol changes occurring in the course of wine-making. *Analytica Chimica Acta*, 458(1), 15-27.
197. Drinkine, J., Lopes, P., Kennedy, J. A., Teissèdre, P. L., & Saucier, C. (2007). Analysis of ethylidene-bridged flavan-3-ols in wine. *Journal of agricultural and food chemistry*, 55(4), 1109-1116.
198. Drinkine, J., Lopes, P., Kennedy, J. A., Teissèdre, P. L., & Saucier, C. (2007). Ethylidene-bridged flavan-3-ols in red wine and correlation with wine age. *Journal of Agricultural and Food Chemistry*, 55(15), 6292-6299.
199. Quaglieri, C., Prieto-Perea, N., Berrueta, L. A., Gallo, B., Rasines-Perea, Z., Jourdes, M., & Teissèdre, P. L. (2017). Comparison of Aquitaine and rioja red wines: Characterization of their phenolic composition and evolution from 2000 to 2013. *Molecules*, 22(2), 192.
200. Abyari, M., Heidari, R., & Jamei, R. (2006). The effect of heating, UV irradiation and PH on stability of Siahe Sardasht Grape anthocyanin–copigment complex. *Journal of Biological Science*, 6(4), 638-645.
201. Jurd, L. (1969). Anthocyanidins and related compounds—XV: The effects of sunlight on flavylum salt-chalcone equilibrium in acid solutions. *Tetrahedron*, 25(11), 2367-2380.

202. Sadilova, E., Carle, R., & Stintzing, F. C. (2007). Thermal degradation of anthocyanins and its impact on color and in vitro antioxidant capacity. *Molecular nutrition & food research*, 51(12), 1461-1471.
203. Hillmann, M. C., Burin, V. M., & Bordignon-Luiz, M. T. (2011). Thermal degradation kinetics of anthocyanins in grape juice and concentrate. *International journal of food science & technology*, 46(9), 1997-2000.
204. Hrazdina, G., & Franzese, A. J. (1974). Oxidation products of acylated anthocyanins under acidic and neutral conditions. *Phytochemistry*, 13(1), 231-234.
205. Kader, F., Irmouli, M., Zitouni, N., Nicolas, J. P., & Metche, M. (1999). Degradation of cyanidin 3-glucoside by caffeic acid o-quinone. Determination of the stoichiometry and characterization of the degradation products. *Journal of Agricultural and Food Chemistry*, 47(11), 4625-4630.
206. Oliveira, C. M., Barros, A. S., Ferreira, A. C. S., & Silva, A. M. (2015). Influence of the temperature and oxygen exposure in red Port wine: A kinetic approach. *Food Research International*, 75, 337-347.
207. Piffaut, B., Kader, F., Girardin, M., & Metche, M. (1994). Comparative degradation pathways of malvidin 3, 5-diglucoside after enzymatic and thermal treatments. *Food Chemistry*, 50(2), 115-120.
208. Trouillas, P., Sancho-García, J. C., De Freitas, V., Gierschner, J., Otyepka, M., & Dangles, O. (2016). Stabilizing and modulating color by copigmentation: Insights from theory and experiment. *Chemical reviews*, 116(9), 4937-4982.
209. Kertesz, M. (2019). Pancake bonding: An unusual Pi-stacking interaction. *Chemistry—A European Journal*, 25(2), 400-416.
210. Scheffeldt, P., & Hrazdina, G. (1978). Co-pigmentation of anthocyanins under physiological conditions. *Journal of Food Science*, 43(2), 517-520.
211. Boulton, R. (2001). The copigmentation of anthocyanins and its role in the color of red wine: A critical review. *American journal of enology and viticulture*, 52(2), 67-87.
212. Cavalcanti, R. N., Santos, D. T., & Meireles, M. A. A. (2011). Non-thermal stabilization mechanisms of anthocyanins in model and food systems—An overview. *Food research international*, 44(2), 499-509.
213. Berké, B., & de Freitas, V. A. (2005). Influence of procyanidin structures on their ability to complex with oenin. *Food chemistry*, 90(3), 453-460.
214. Davies, A. J., & Mazza, G. (1993). Copigmentation of simple and acylated anthocyanins with colorless phenolic compounds. *Journal of Agricultural and Food Chemistry*, 41(5), 716-720.
215. Brouillard, R., Wigand, M. C., Dangles, O., & Cheminat, A. (1991). pH and solvent effects on the copigmentation reaction of malvin with polyphenols, purine and pyrimidine derivatives. *Journal of the Chemical Society, Perkin Transactions 2*, (8), 1235-1241.

216. Mazza, G., & Brouillard, R. (1990). The mechanism of co-pigmentation of anthocyanins in aqueous solutions. *Phytochemistry*, 29(4), 1097-1102.
217. Teixeira, N., Cruz, L., Brás, N. F., Mateus, N., Ramos, M. J., & de Freitas, V. (2013). Structural features of copigmentation of oenin with different polyphenol copigments. *Journal of agricultural and food chemistry*, 61(28), 6942-6948.
218. Zhang, B., Wang, Q., Zhou, P. P., Li, N. N., & Han, S. Y. (2020). Copigmentation evidence of oenin with phenolic compounds: A comparative study of spectrographic, thermodynamic and theoretical data. *Food chemistry*, 313, 126163.
219. Bakker, J., Bridle, P., Honda, T., Kuwano, H., Saito, N., Terahara, N., & Timberlake, C. F. (1997). Identification of an anthocyanin occurring in some red wines. *Phytochemistry*, 44(7), 1375-1382.
220. Bakker, J., & Timberlake, C. F. (1985). The distribution and content of anthocyanins in young port wines as determined by high performance liquid chromatography. *Journal of the Science of Food and Agriculture*, 36(12), 1325-1333.
221. Bakker, J., & Timberlake, C. F. (1997). Isolation, identification, and characterization of new color-stable anthocyanins occurring in some red wines. *Journal of Agricultural and Food Chemistry*, 45(1), 35-43.
222. de Freitas, V., & Mateus, N. (2011). Formation of pyranoanthocyanins in red wines: a new and diverse class of anthocyanin derivatives. *Analytical and bioanalytical chemistry*, 401(5), 1463-1473.
223. Mateus, N., Silva, A. M., Santos-Buelga, C., Rivas-Gonzalo, J. C., & de Freitas, V. (2002). Identification of anthocyanin-flavanol pigments in red wines by NMR and mass spectrometry. *Journal of Agricultural and Food Chemistry*, 50(7), 2110-2116.
224. Quagliari, C., Jourdes, M., Waffo-Teguo, P., & Teissedre, P. L. (2017). Updated knowledge about pyranoanthocyanins: Impact of oxygen on their contents, and contribution in the winemaking process to overall wine color. *Trends in Food Science & Technology*, 67, 139-149.
225. Rentzsch, M., Schwarz, M., Winterhalter, P., Blanco-Vega, D., & Hermosín-Gutiérrez, I. (2010). Survey on the content of vitisin A and hydroxyphenyl-pyranoanthocyanins in Tempranillo wines. *Food Chemistry*, 119(4), 1426-1434.
226. Blanco-Vega, D., Gómez-Alonso, S., & Hermosín-Gutiérrez, I. (2014). Identification, content and distribution of anthocyanins and low molecular weight anthocyanin-derived pigments in Spanish commercial red wines. *Food chemistry*, 158, 449-458.
227. Rentzsch, M., Weber, F., Durner, D., Fischer, U., & Winterhalter, P. (2009). Variation of pyranoanthocyanins in red wines of different varieties and vintages and the impact of pinotin A addition on their color parameters. *European Food Research and Technology*, 229(4), 689-696.

228. Hagerman, A. E. (1989). Chemistry of tannin-protein complexation. In *Chemistry and significance of condensed tannins* (pp. 323-333). Springer, Boston, MA.
229. Haslam, E., Lilley, T. H., & Butler, L. G. (1988). Natural astringency in foodstuffs—a molecular interpretation. *Critical Reviews in Food Science & Nutrition*, 27(1), 1-40.
230. Waters, E. J., Shirley, N. J., & Williams, P. J. (1996). Nuisance proteins of wine are grape pathogenesis-related proteins. *Journal of Agricultural and Food Chemistry*, 44(1), 3-5.
231. Yokotsuka, K., Nozaki, K., & Takayanagi, T. (1994). Characterization of soluble glycoproteins in red wine. *American journal of enology and viticulture*, 45(4), 410-416.
232. Maury, C., Sarni-Manchado, P., Lefebvre, S., Cheynier, V., & Moutounet, M. (2003). Influence of fining with plant proteins on proanthocyanidin composition of red wines. *American Journal of Enology and Viticulture*, 54(2), 105-111.
233. Smith, M. R., Penner, M. H., Bennett, S. E., & Bakalinsky, A. T. (2011). Quantitative colorimetric assay for total protein applied to the red wine pinot noir. *Journal of agricultural and food chemistry*, 59(13), 6871-6876.
234. Maury, C., Sarni-Manchado, P., Lefebvre, S., Cheynier, V., & Moutounet, M. (2001). Influence of fining with different molecular weight gelatins on proanthocyanidin composition and perception of wines. *American journal of enology and viticulture*, 52(2), 140-145.
235. Ricardo-da-Silva, J. M., Cheynier, V., Souquet, J. M., Moutounet, M., Cabanis, J. C., & Bourzeix, M. (1991). Interaction of grape seed procyanidins with various proteins in relation to wine fining. *Journal of the Science of Food and Agriculture*, 57(1), 111-125.
236. Guadalupe, Z., Martínez-Pinilla, O., Garrido, Á., Carrillo, J. D., & Ayestarán, B. (2012). Quantitative determination of wine polysaccharides by gas chromatography–mass spectrometry (GC–MS) and size exclusion chromatography (SEC). *Food Chemistry*, 131(1), 367-374.
237. Alcalde-Eon, C., Garcia-Estevez, I., Puente, V., Rivas-Gonzalo, J. C., & Escribano-Bailón, M. T. (2014). Color stabilization of red wines. A chemical and colloidal approach. *Journal of agricultural and food chemistry*, 62(29), 6984-6994.
238. Carvalho, E., Mateus, N., Plet, B., Pianet, I., Dufourc, E., & De Freitas, V. (2006). Influence of wine pectic polysaccharides on the interactions between condensed tannins and salivary proteins. *Journal of Agricultural and Food Chemistry*, 54(23), 8936-8944.
239. Maury, C., Sarni-Manchado, P., Poinssaut, P., Cheynier, V., & Moutounet, M. (2016). Influence of polysaccharides and glycerol on proanthocyanidin precipitation by protein fining agents. *Food Hydrocolloids*, 60, 598-605.
240. Jaeckels, N., Meier, M., Dietrich, H., Will, F., Decker, H., & Fronk, P. (2016). Influence of polysaccharides on wine protein aggregation. *Food chemistry*, 200, 38-45.

241. Li, S., Wilkinson, K. L., Mierczynska-Vasilev, A., & Bindon, K. A. (2019). Applying nanoparticle tracking analysis to characterize the polydispersity of aggregates resulting from tannin–polysaccharide interactions in wine-like media. *Molecules*, 24(11), 2100.
242. Pokorný, J., Filipů, M., & Pudil, F. (1998). Prediction of odour and flavour acceptancies of white wines on the basis of their colour. *Food/Nahrung*, 42(06), 412-415.
243. Spence, C. (2010). The color of wine–Part 1. *The World of Fine Wine*, 28(2010), 122-129.
244. Parr, W. V., Geoffrey White, K., & Heatherbell, D. A. (2003). The nose knows: Influence of colour on perception of wine aroma. *Journal of wine research*, 14(2-3), 79-101.
245. Brossaud, F., Cheynier, V., & Noble, A. C. (2001). Bitterness and astringency of grape and wine polyphenols. *Australian Journal of grape and wine research*, 7(1), 33-39.
246. Ma, W., Guo, A., Zhang, Y., Wang, H., Liu, Y., & Li, H. (2014). A review on astringency and bitterness perception of tannins in wine. *Trends in Food Science & Technology*, 40(1), 6-19.
247. American Society for Testing and Materials. (1978). Standard definitions of terms relating to sensory evaluation of materials and products. Astm.
248. Vidal, L., Giménez, A., Medina, K., Boido, E., & Ares, G. (2015). How do consumers describe wine astringency?. *Food Research International*, 78, 321-326.
249. Bate-Smith, E. C. (1954). Astringency in foods. *Food*, 23(124).
250. Breslin, P. A. (1996). Interactions among salty, sour and bitter compounds. *Trends in Food Science & Technology*, 7(12), 390-399.
251. Guinard, J. X., Pangborn, R. M., & Lewis, M. J. (1986). The time-course of astringency in wine upon repeated ingestion. *American Journal of Enology and Viticulture*, 37(3), 184-189.
252. Hatano, T., & Hemingway, R. W. (1996). Association of (+)-catechin and catechin-(4 $\alpha$ → 8)-catechin with oligopeptides. *Chemical Communications*, (22), 2537-2538.
253. Jöbstl, E., Howse, J. R., Fairclough, J. P. A., & Williamson, M. P. (2006). Noncovalent cross-linking of casein by epigallocatechin gallate characterized by single molecule force microscopy. *Journal of Agricultural and Food Chemistry*, 54(12), 4077-4081.
254. Oh, H. I., Hoff, J. E., Armstrong, G. S., & Haff, L. A. (1980). Hydrophobic interaction in tannin-protein complexes. *Journal of Agricultural and Food Chemistry*, 28(2), 394-398.
255. Wróblewski, K., Muhandiram, R., Chakrabarty, A., & Bennick, A. (2001). The molecular interaction of human salivary histatins with polyphenolic compounds. *European Journal of Biochemistry*, 268(16), 4384-4397.
256. Jaggupilli, A., Howard, R., Upadhyaya, J. D., Bhullar, R. P., & Chelikani, P. (2016). Bitter taste receptors: Novel insights into the biochemistry and pharmacology. *The international journal of biochemistry & cell biology*, 77, 184-196.

257. Peng, Y., Gillis-Smith, S., Jin, H., Tränkner, D., Ryba, N. J., & Zuker, C. S. (2015). Sweet and bitter taste in the brain of awake behaving animals. *Nature*, *527*(7579), 512-515.
258. Roudnitzky, N., Behrens, M., Engel, A., Kohl, S., Thalmann, S., Hübner, S., & Meyerhof, W. (2015). Receptor polymorphism and genomic structure interact to shape bitter taste perception. *PLoS Genet*, *11*(9)
259. Adler, E., Hoon, M. A., Mueller, K. L., Chandrashekar, J., Ryba, N. J., & Zuker, C. S. (2000). A novel family of mammalian taste receptors. *Cell*, *100*(6), 693-702.
260. Luo, Y., Kong, L., Xue, R., Wang, W., & Xia, X. (2020). Bitterness in alcoholic beverages: The profiles of perception, constituents, and contributors. *Trends in Food Science & Technology*, *96*, 222-232.
261. Deshpande, D. A., Wang, W. C., McIlmoyle, E. L., Robinett, K. S., Schillinger, R. M., An, S. S., ... & Liggett, S. B. (2010). Bitter taste receptors on airway smooth muscle bronchodilate by localized calcium signaling and reverse obstruction. *Nature medicine*, *16*(11), 1299.
262. Roper, S. D., & Chaudhari, N. (2017). Taste buds: cells, signals and synapses. *Nature Reviews Neuroscience*, *18*(8), 485-497.
263. Arnold, R. A., Noble, A. C., & Singleton, V. L. (1980). Bitterness and astringency of phenolic fractions in wine. *Journal of agricultural and food chemistry*, *28*(3), 675-678.
264. Guinard, J. X., Pangborn, R. M., & Lewis, M. J. (1986). Preliminary studies on acidity-astringency interactions in model solutions and wines. *Journal of the Science of Food and Agriculture*, *37*(8), 811-817.
265. Lea, A. G. (1992). Flavor, color, and stability in fruit products: The effect of polyphenols. In *Plant polyphenols* (pp. 827-847). Springer, Boston, MA.
266. Ma, W., Waffo-Teguo, P., Jourdes, M., Li, H., & Teissedre, P. L. (2016). Chemical affinity between tannin size and salivary protein binding abilities: Implications for wine astringency. *PloS one*, *11*(8), e0161095.
267. Kallithraka, S., Bakker, J., & Clifford, M. N. (1997). Evaluation of bitterness and astringency of (+)-catechin and (-)-epicatechin in red wine and in model solution. *Journal of Sensory Studies*, *12*(1), 25-37.
268. Robichaud, J. L., & Noble, A. C. (1990). Astringency and bitterness of selected phenolics in wine. *Journal of the Science of Food and Agriculture*, *53*(3), 343-353.
269. Chira, K., Zeng, L., Le Floch, A., Péchamat, L., Jourdes, M., & Teissedre, P. L. (2015). Compositional and sensory characterization of grape proanthocyanidins and oak wood ellagitannin. *Tetrahedron*, *71*(20), 2999-3006.
270. Ferrer-Gallego, R., García-Marino, M., Hernández-Hierro, J. M., Rivas-Gonzalo, J. C., & Escribano-Bailón, M. T. (2010). Statistical correlation between flavanolic composition, colour and sensorial parameters in grape seed during ripening. *Analytica Chimica Acta*, *660*(1-2), 22-28.

271. Mirabel, M. (2000). *Caractéristiques chimiques et organoleptiques des tanins des raisins de Vitis vinifera var. Merlot et Cabernet Sauvignon issus de différents terroirs bordelais* (Doctoral dissertation, Bordeaux 2).
272. Fontoin, H., Saucier, C., Teissedre, P. L., & Glories, Y. (2008). Effect of pH, ethanol and acidity on astringency and bitterness of grape seed tannin oligomers in model wine solution. *Food Quality and Preference*, 19(3), 286-291.
273. Pissoni, M. A., Waffo-Teguo, P., Ma, W., Jourdes, M., Rolle, L., & Teissedre, P. L. (2018). Chemical and sensorial investigation of in-mouth sensory properties of grape anthocyanins. *Scientific reports*, 8(1), 1-13.
274. Ribéreau-Gayon, P., & Stonestreet, E. (1965). Determination of anthocyanins in red wine. *Bulletin de la Societe chimique de France*, 9, 2649-2652.
275. Harbertson, J. F., Picciotto, E. A., & Adams, D. O. (2003). Measurement of polymeric pigments in grape berry extract and wines using a protein precipitation assay combined with bisulfite bleaching. *American Journal of Enology and Viticulture*, 54(4), 301-306.
276. Baldi, A., Romani, A., Mulinacci, N., Vincieri, F. F., & Casetta, B. (1995). HPLC/MS application to anthocyanins of *Vitis vinifera* L. *Journal of Agricultural and Food Chemistry*, 43(8), 2104-2109.
277. Kennedy, J. A., & Jones, G. P. (2001). Analysis of proanthocyanidin cleavage products following acid-catalysis in the presence of excess phloroglucinol. *Journal of Agricultural and Food Chemistry*, 49(4), 1740-1746.
278. Swinehart, D. F. (1962). The beer-lambert law. *Journal of chemical education*, 39(7), 333.
279. Cozzolino, D., Cynkar, W. U., Shah, N., Damberg, R. G., & Smith, P. A. (2009). A brief introduction to multivariate methods in grape and wine analysis. *International Journal of Wine Research*, 1, 123-130.
280. Giusti, M. M., & Wrolstad, R. E. (2001). Characterization and measurement of anthocyanins by UV-visible spectroscopy. *Current protocols in food analytical chemistry*, (1), F1-2.
281. Ribéreau-Gayon, P., Glories, Y., Maujean, A., & Dubourdieu, D. (2012). *Traité d'oenologie-Tome 2-6e éd.-Chimie du vin. Stabilisation et traitements*. Dunod.
282. Boulet, J. C., Ducasse, M. A., & Cheynier, V. (2017). Ultraviolet spectroscopy study of phenolic substances and other major compounds in red wines: Relationship between astringency and the concentration of phenolic substances. *Australian journal of grape and wine research*, 23(2), 193-199.
283. Alexandre-Tudo, J. L., & Du Toit, W. (2018). The role of UV-visible spectroscopy for phenolic compounds quantification in winemaking. *Frontiers and new trends in the science of fermented food and beverages*.
284. Doyle, W. M. (1992). Principles and applications of Fourier transform infrared (FTIR) process analysis. *Process Control Qual*, 2(1), 11-41.

285. Cohen, L. (1989). Time-frequency distributions-a review. *Proceedings of the IEEE*, 77(7), 941-981.
286. Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, 36(5), 961-1005.
287. Aleixandre-Tudo, J. L., Nieuwoudt, H., Aleixandre, J. L., & du Toit, W. (2018). Chemometric compositional analysis of phenolic compounds in fermenting samples and wines using different infrared spectroscopy techniques. *Talanta*, 176, 526-536.
288. Fernández, K., & Agosin, E. (2007). Quantitative analysis of red wine tannins using Fourier-transform mid-infrared spectrometry. *Journal of Agricultural and Food Chemistry*, 55(18), 7294-7300.
289. 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 spectroscopy. *Italian journal of food science*, 18(4).
290. Abbas, O., Compère, G., Larondelle, Y., Pompeu, D., Rogez, H., & Baeten, V. (2017). Phenolic compound explorer: a mid-infrared spectroscopy database. *Vibrational Spectroscopy*, 92, 111-118.
291. Edelmann, A., & Lendl, B. (2002). Toward the optical tongue: Flow-through sensing of tannin– protein interactions based on FTIR spectroscopy. *Journal of the American Chemical Society*, 124(49), 14741-14747.
292. Shurvell, H. F. (2006). Spectra–structure correlations in the mid-and far-infrared. *Handbook of vibrational spectroscopy*.
293. Foo, L. Y. (1981). Proanthocyanidins: Gross chemical structures by infrared spectra. *Phytochemistry*, 20(6), 1397-1402.
294. Cozzolino, D., Damberg, R. G., Janik, L., Cynkar, W. U., & Gishen, M. (2006). Analysis of grapes and wine by near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, 14(5), 279-289.
295. McClure, W. F. (2003). 204 years of near infrared technology: 1800–2003. *Journal of Near Infrared Spectroscopy*, 11(6), 487-518.
296. Ziegel, E. R. (2004). A user-friendly guide to multivariate calibration and classification.
297. Basalekou, M., Pappas, C., Kotseridis, Y., Tarantilis, P. A., Kontaxakis, E., & Kallithraka, S. (2017). Red wine age estimation by the alteration of its color parameters: Fourier transform infrared spectroscopy as a tool to monitor wine maturation time. *Journal of analytical methods in chemistry*, 2017.
298. Bauer, R., Nieuwoudt, H., Bauer, F. F., Kossmann, J., Koch, K. R., & Esbensen, K. H. (2008). FTIR spectroscopy for grape and wine analysis.

299. Moreira, J. L., & Santos, L. (2005). Analysis of organic acids in wines by Fourier-transform infrared spectroscopy. *Analytical and bioanalytical chemistry*, 382(2), 421-425.
300. Pizarro, C., González-Sáiz, J. M., Esteban-Díez, I., & Orío, P. (2011). Prediction of total and volatile acidity in red wines by Fourier-transform mid-infrared spectroscopy and iterative predictor weighting. *Analytical and bioanalytical chemistry*, 399(6), 2061-2072.
301. Chin, W. W. (1998). Commentary: Issues and opinion on structural equation modeling.
302. Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational statistics & data analysis*, 48(1), 159-205.
303. Geladi, P. (2003). Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 58(5), 767-782.
304. Wold, S. (1995). Chemometrics; what do we mean with it, and what do we want from it?. *Chemometrics and Intelligent Laboratory Systems*, 30(1), 109-115.
305. Munck, L., Nørgaard, L., Engelsen, S. B., Bro, R., & Andersson, C. A. (1998). Chemometrics in food science—a demonstration of the feasibility of a highly exploratory, inductive evaluation strategy of fundamental scientific significance. *Chemometrics and Intelligent Laboratory Systems*, 44(1-2), 31-60.
306. Baardseth, P., Helgesen, H., & Isaksson, T. (1996). Multivariate techniques in the analysis of meat quality. *Meat Science*, 43, 135-149.
307. Guinot, C., Latreille, J., & Tenenhaus, M. (2001). PLS path modelling and multiple table analysis. Application to the cosmetic habits of women in Ile-de-France. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 247-259.
308. Pagès, J., & Tenenhaus, M. (2001). Multiple factor analysis combined with PLS path modelling. Application to the analysis of relationships between physicochemical variables, sensory profiles and hedonic judgements. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 261-273.
309. Brereton, R. G. (2000). Introduction to multivariate calibration in analytical chemistry. *Analyst*, 125(11), 2125-2154.
310. Xu, L., Zhou, Y. P., Tang, L. J., Wu, H. L., Jiang, J. H., Shen, G. L., & Yu, R. Q. (2008). Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. *Analytica chimica acta*, 616(2), 138-143.
311. Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2016). Mean centering helps alleviate “micro” but not “macro” multicollinearity. *Behavior research methods*, 48(4), 1308-1317.
312. Nicolai, B. M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K. I., & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. *Postharvest biology and technology*, 46(2), 99-118.

313. Murray, I. (1999). NIR spectroscopy of food: simple things, subtle things and spectra. *NIR news*, 10(2), 10-12.
314. Roberts, C. A., Workman, J., & Reeves, J. B. (Eds.). (2004). *Near-infrared spectroscopy in agriculture* (Vol. 44). Madison: American Society of Agronomy.
315. Murray, I. (1993). Forage analysis by near infrared spectroscopy. *Sward measurement handbook*, 285-312.
316. Aleixandre-Tudo, J. L., Buica, A., Nieuwoudt, H., Aleixandre, J. L., & du Toit, W. (2017). Spectrophotometric analysis of phenolic compounds in grapes and wines. *Journal of agricultural and food chemistry*, 65(20), 4009-4026.
317. Dardenne, P. (2010). Some considerations about NIR spectroscopy: Closing speech at NIR-2009. *NIR news*, 21(1), 8-14.
318. Galvaoa, R. K. H., Araujob, M. C. U., Joséb, G. E., Pontesb, M. J. C., & Silvab, E. C. (2005). A method for calibration and validation subset partitioning. *Talanta*, 67(4), 736-740.
319. Zeaiter, M., Roger, J. M., Bellon-Maurel, V., & Rutledge, D. N. (2004). Robustness of models developed by multivariate calibration. Part I: The assessment of robustness. *TrAC Trends in Analytical Chemistry*, 23(2), 157-170.
320. Gordon, R., Chapman, J., Power, A., Chandra, S., Roberts, J., & Cozzolino, D. (2019). Mid-infrared spectroscopy coupled with chemometrics to identify spectral variability in Australian barley samples from different production regions. *Journal of Cereal Science*, 85, 41-47.
321. Martens, H., & Naes, T. (1992). *Multivariate calibration*. John Wiley & Sons.
322. Kyrleou, M., Pappas, C., Voskidi, E., Kotseridis, Y., Basalekou, M., Tarantilis, P. A., & Kallithraka, S. (2015). Diffuse reflectance Fourier transform infrared spectroscopy for simultaneous quantification of total phenolics and condensed tannins contained in grape seeds. *Industrial Crops and Products*, 74, 784-791.
323. Janik, L. J., Cozzolino, D., Damberg, R., Cynkar, W., & Gishen, M. (2007). The prediction of total anthocyanin concentration in red-grape homogenates using visible-near-infrared spectroscopy and artificial neural networks. *Analytica chimica acta*, 594(1), 107-118.
324. García-Jares, C., & Medina, B. (1995). Prediction of some physico-chemical parameters in red wines from ultraviolet-visible spectra using a partial least-squares model in latent variables. *Analyst*, 120(7), 1891-1896.
325. Boulet, J. C., Trarieux, C., Souquet, J. M., Ducasse, M. A., Caillé, S., Samson, A., ... & Cheynier, V. (2016). Models based on ultraviolet spectroscopy, polyphenols, oligosaccharides and polysaccharides for prediction of wine astringency. *Food chemistry*, 190, 357-363.
326. Damberg, R. G., Mercurio, M. D., Kassara, S., Cozzolino, D., & Smith, P. A. (2012). Rapid measurement of methyl cellulose precipitable tannins using ultraviolet spectroscopy with chemometrics: Application to red wine and inter-laboratory calibration transfer. *Applied Spectroscopy*, 66(6), 656-664.

327. Mercurio, M. D., Damberg, R. G., Herderich, M. J., & Smith, P. A. (2007). High throughput analysis of red wine and grape phenolics adaptation and validation of methyl cellulose precipitable tannin assay and modified somers color assay to a rapid 96 well plate format. *Journal of agricultural and food chemistry*, 55(12), 4651-4657.
328. Di Egidio, V., Sinelli, N., Giovanelli, G., Moles, A., & Casiraghi, E. (2010). NIR and MIR spectroscopy as rapid methods to monitor red wine fermentation. *European Food Research and Technology*, 230(6), 947-955.
329. Rasines-Perea, Z., Prieto-Perea, N., Romera-Fernández, M., Berrueta, L. A., & Gallo, B. (2015). Fast determination of anthocyanins in red grape musts by Fourier transform mid-infrared spectroscopy and partial least squares regression. *European Food Research and Technology*, 240(5), 897-908.
330. Basalekou, M., Kallithraka, S., Tarantilis, P. A., Kotseridis, Y., & Pappas, C. (2019). Ellagitannins in wines: Future prospects in methods of analysis using FT-IR spectroscopy. *LWT*, 101, 48-53.
331. Glories, Y. (1993). Maturité phénolique du raisin, conséquences technologiques: application aux millésimes 1991 et 1992. In *Journée technique du CIVB: Actes du colloque* (pp. 56-61).
332. Kontoudakis, N., Esteruelas, M., Fort, F., Canals, J. M., & Zamora, F. (2010). Comparison of methods for estimating phenolic maturity in grapes: Correlation between predicted and obtained parameters. *Analytica chimica acta*, 660(1-2), 127-133.
333. Avizcuri, J. M., Sáenz-Navajas, M. P., Echávarri, J. F., Ferreira, V., & Fernández-Zurbano, P. (2016). Evaluation of the impact of initial red wine composition on changes in color and anthocyanin content during bottle storage. *Food chemistry*, 213, 123-134.
334. Picariello, L., Gambuti, A., Picariello, B., & Moio, L. (2017). Evolution of pigments, tannins and acetaldehyde during forced oxidation of red wine: Effect of tannins addition. *LWT*, 77, 370-375.
335. Adams, D. O., & Harbertson, J. F. (1999). Use of alkaline phosphatase for the analysis of tannins in grapes and red wines. *American journal of enology and viticulture*, 50(3), 247-252.
336. Harbertson, J. F., Kilmister, R. L., Kelm, M. A., & Downey, M. O. (2014). Impact of condensed tannin size as individual and mixed polymers on bovine serum albumin precipitation. *Food chemistry*, 160, 16-21.
337. Sarneckis, C. J., Damberg, R. G., Jones, P., Mercurio, M., Herderich, M. J., & Smith, P. A. (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(1), 39-49.
338. Chira, K. (2009). *Structures moléculaire et perception tannique des raisins et des vins (Cabernet-Sauvignon, Merlot) du Bordelais* (Doctoral dissertation, Bordeaux 2).
339. De Jong, S. (1993). SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and intelligent laboratory systems*, 18(3), 251-263.

340. Kennard, R. W., & Stone, L. A. (1969). Computer aided design of experiments. *Technometrics*, *11*(1), 137-148.
341. Sessa, C., Vila, A., & García, J. F. (2011). Determination of detection limits for SEM-EDS and m-FTIR analysis of artwork. *Analytical and bioanalytical chemistry*, *400*(7), 2241-2251.
342. Levengood, J., & Boulton, R. (2004). The variation in the color due to copigmentation in young Cabernet Sauvignon wines.
343. González-Neves, G., Barreiro, L., Gil, G., Franco, J., Ferrer, M., Moutounet, M., & Carbonneau, A. (2004). Anthocyanic composition of Tannat grapes from the south region of Uruguay. *Analytica Chimica Acta*, *513*(1), 197-202.
344. Vian, M. A., Tomao, V., Coulomb, P. O., Lacombe, J. M., & Dangles, O. (2006). Comparison of the anthocyanin composition during ripening of Syrah grapes grown using organic or conventional agricultural practices. *Journal of Agricultural and Food Chemistry*, *54*(15), 5230-5235.
345. Marquez, A., Dueñas, M., Serratos, M. P., & Merida, J. (2012). Formation of vitisins and anthocyanin–flavanol adducts during red grape drying. *Journal of agricultural and food chemistry*, *60*(27), 6866-6874.
346. Vidal, S., Meudec, E., Cheynier, V., Skouroumounis, G., & Hayasaka, Y. (2004). Mass spectrometric evidence for the existence of oligomeric anthocyanins in grape skins. *Journal of Agricultural and Food Chemistry*, *52*(23), 7144-7151.
347. Fournand, D., Vicens, A., Sidhoum, L., Souquet, J. M., Moutounet, M., & Cheynier, V. (2006). Accumulation and extractability of grape skin tannins and anthocyanins at different advanced physiological stages. *Journal of Agricultural and Food Chemistry*, *54*(19), 7331-7338.
348. Hanlin, R. L., & Downey, M. O. (2009). Condensed tannin accumulation and composition in skin of Shiraz and Cabernet Sauvignon grapes during berry development. *American Journal of Enology and Viticulture*, *60*(1), 13-23.
349. Rousserie, P., Lacampagne, S., Vanbrabant, S., Rabot, A., & Geny-Denis, L. (2020). Influence of berry ripeness on seed tannins extraction in wine. *Food chemistry*, *315*, 126307.
350. Canals, R., Llaudy, M. C., Valls, J., Canals, J. M., & Zamora, F. (2005). Influence of ethanol concentration on the extraction of color and phenolic compounds from the skin and seeds of Tempranillo grapes at different stages of ripening. *Journal of Agricultural and Food Chemistry*, *53*(10), 4019-4025.
351. de Freitas, V. A., & Mateus, N. (2010). Updating wine pigments. *Recent advances in polyphenol research, Volume 2*, 59-80.
352. Iacobucci, G. A., & Sweeny, J. G. (1983). The chemistry of anthocyanins, anthocyanidins and related flavylum salts. *Tetrahedron*, *39*(19), 3005-3038.

353. Petrozziello, M., Torchio, F., Piano, F., Giacosa, S., Ugliano, M., Bosso, A., & Rolle, L. (2018). Impact of increasing levels of oxygen consumption on the evolution of color, phenolic, and volatile compounds of Nebbiolo wines. *Frontiers in chemistry*, 6, 137.
354. Ferrer-Gallego, R., Hernández-Hierro, J. M., Rivas-Gonzalo, J. C., & Escribano-Bailón, M. T. (2011). Determination of phenolic compounds of grape skins during ripening by NIR spectroscopy. *LWT-Food Science and Technology*, 44(4), 847-853.
355. Michel, J., Albertin, W., Jourdes, M., Le Floch, A., Giordanengo, T., Mourey, N., & Teissedre, P. L. (2016). Variations in oxygen and ellagitannins, and organoleptic properties of red wine aged in French oak barrels classified by a near infrared system. *Food chemistry*, 204, 381-390.
356. Moutounet, M., Mazauric, J. P., Saint Pierre, B., & Hanocq, J. F. (1998). Gaseous exchange in wines stored in barrels. *Journal des Sciences et Techniques de la Tonnelerie (France)*.
357. Sánchez-Gómez, R., del Alamo-Sanza, M., Martínez-Martínez, V., & Nevares, I. (2020). Study of the role of oxygen in the evolution of red wine colour under different ageing conditions in barrels and bottles. *Food Chemistry*, 328, 127040.

