

APPRÉCIATION ET COMPARAISON « EN LIGNE » DE LA COULEUR ET DE LA STRUCTURE PHÉNOLIQUE DES VINS ROUGES

APPRECIATION AND «ON-LINE» COMPARISON OF COLOUR AND PHENOLIC STRUCTURE OF RED WINES

J. BLOUIN¹, D. PIERRE-LOTI-VIAUD²,
Nathalie CRACHEREAU³ et J.-P. RICHARD⁴

1 : Chambre d'Agriculture de la Gironde,
39, rue Michel Montaigne 33490 Blanquefort, France

2 : LSTA Laboratoire de statistique théorique et appliquée Université Paris VI
6, place Jussieu, 75005 Paris, France

3 : Fédération des centres œnologiques de la Gironde, 39 rue Michel Montaigne,
33490 Blanquefort, France

4 : Société ABC, 7, rue Chevert, 75007 Paris, France

Résumé : L'appréciation visuelle de la couleur des vins reste décisive mais sa caractérisation objective est essentielle. Depuis plus d'un siècle, de nombreuses méthodes ont été proposées. La technique que nous décrivons ici est nouvelle sous deux aspects principaux :

- d'une part, elle permet un traitement quantitatif de l'ensemble du spectre, incluant à la fois les pics majeurs, classiques, et la partie résiduelle (partie du spectre restant après élimination des pics majeurs), inexploitable jusqu'à présent. On décompose ainsi le spectre en un vecteur dont le nombre de composantes scalaires est calculé de telle sorte que la perte d'information résultant de cette décomposition soit fixée à l'avance ;

- d'autre part, le vecteur représentant la totalité du spectre est utilisé de façon originale pour des classifications et des comparaisons multivariées.

Nous avons appliqué cette technique à la classification de 213 vins connus de la région bordelaise et comparé nos résultats aux méthodes habituelles.

Notre étude montre que tout vin peut être défini désormais par un vecteur dont les composantes peuvent être mixtes: spectrales, chromatographiques, analytiques, organoleptiques. Parmi les nombreuses composantes disponibles on ne retient que les plus discriminantes pour ainsi améliorer les classifications et les comparaisons.

Certaines de ces composantes sont directement liées à la concentration en divers constituants du vin. La rapidité des mesures photométriques et de leur traitement informatique permet de parler de résultats en ligne.

Summary: If, on one hand, the visual appreciation of wines color is of the greatest importance, on the other hand, its objective characterisation is essential. For more than a century, many methods have been proposed. The technique we describe here is new under two main aspects:

- first, it allows a quantitative treatment of the whole spectrum, including both major, classical peaks and the residual part (part of the spectrum remaining after elimination of major peaks), unexploitable until now. The spectrum is decomposed in a vector whose the number of scalars components is computed in such a way that the loss of information resulting of this decomposition is determined in advance;

- secondly, the vector representing the whole spectrum is used in an original way for classifications and multivariate comparisons.

We applied this technique to classification of 213 identified wines of Bordeaux area and compared our results to usual methods. Beyond this application, one can see that any wine can be defined, from now, by a vector whose components are of any kind: spectral, chromatographic, analytic, organoleptic. The great number of possible components permits us to eliminate the less discriminant ones and to select, iteratively, the more discriminant ones and to improve, by the way, classifications and comparisons. The measurements and their computerized treatment are so fast that we can speak of on-line results.

Mots clés : vins rouges, couleur, composés phénoliques, empreintes spectrales.

Key words: red wines, colour, phenolic compounds, spectral fingerprints.

INTRODUCTION

La couleur des vins rouges et rosés est un élément très important de leur caractérisation et de leur appréciation par le consommateur. L'appréciation visuelle

reste décisive, mais les mesures physiques sont pratiquées depuis longtemps. Après l'utilisation des colorimètres visuels comparant les vins à des disques de satin colorés (Vino-colorimètre SALLERON-1885), ou à des solutions de référence (ROOS, 1930) sont

apparues les mesures de densité optique (FAURE et PALLU, 1935) qui se généralisèrent avec la multiplication des appareils et la « méthode SUDRAUD » (1958) définissant deux indices (intensité colorante et teinte), complétés par GLORIES (1984). La mesure d'absorption de la lumière ultraviolette (FLANZY et POUX, 1958 ; RIBÉREAU-GAYON, 1970) définit un Indice de Polyphénols Totaux (IPT).

Alors que le vin absorbe la lumière entre les longueurs d'onde de 190 à 820 nm, environ, on n'exploite de façon usuelle que quatre valeurs mesurées à 280, 420, 520 et 620 nm. Ce travail vise à utiliser l'ensemble du spectre d'absorption de la lumière ultraviolette et visible en comparant ces résultats avec diverses méthodes de mesure de la couleur afin de dégager les mesures les plus significatives. A l'aide de méthodes nouvelles de traitement du signal – ici le spectre UV-visible des vins – et de comparaisons des données, nous avons également dégagé une méthodologie d'interprétation de ces données analytiques.

MATÉRIELS ET MÉTHODES

I - OBTENTION D'UNE BASE DE DONNÉES

Nous avons réuni 213 vins issus de divers vignobles bordelais, soit de cépages purs (Merlot, Cabernet Sauvignon, Cabernet franc), soit de leurs assemblages, et de différents millésimes. A l'aide d'un spectrophotomètre, nous avons mesuré les absorbances [] de 190 à 820 nm (lumière transmise), après dilution si nécessaire. L'ensemble de ces vins, rouges, rosés, claires couvrent une très vaste gamme de couleur (intensité colorante entre 1 et 17) et de richesse phénolique (IPT entre 10 et 65).

II - MÉTHODES UTILISÉES POUR LA MESURE DE LA COULEUR DES VINS

1) Méthodes traditionnelles

Rappelons les définitions des indices de couleur permettant d'apprécier la couleur d'un vin :

- a) L'intensité colorante
(somme des absorbances à 420 et 520 nm)

$$IC = A_{420} + A_{520}$$

- b) L'intensité colorante modifiée
(somme des absorbances à 420, 520 et 620 nm)

$$ICM = A_{420} + A_{520} + A_{620}$$

- c) La teinte
(rapport des absorbances à 420 et 520 nm)

$$T = A_{420} / A_{520}$$

- d) La contribution des anthocyanes rouges :

$$DA\% = [1 - (A_{420} + A_{620}) / (2 * A_{520})] * 100$$

- e) Le nombre de couleur. Cette valeur (ROUBERT, 1977), peu usitée, est assimilable à une quantité de couleur.

$$NbC = 300 / (1.280 / A_{625} + 0.735 / A_{650} + 0.39 / A_{675})$$

- 2) Caractéristiques chromatiques des vins : approche normative

Nous avons utilisé les méthodes officielles européennes (Journal Officiel CEE, 1990) et de la Commission Internationale de l'Eclairage (AFNOR, 1984). Nous avons développé les outils de calcul rapide exploitant directement les fichiers de données brutes fournis par le spectrophotomètre.

La norme de la mesure des couleurs, définit une échelle de couleurs dites « uniformes ».

Chaque couleur est représentée par trois mesures :

- la clarté L^* , comprise entre $L^*=0$ (noir) et $L^*=100$ (blanc) ;
- la composante chromatique rouge-vert a^* , allant du vert (-100) au rouge (+100) ;
- la composante chromatique jaune-bleu b^* , allant du bleu (-100) au jaune (+100).

Cette échelle, basée sur la vision des couleurs antagonistes, est appelée « échelle CIE 1976 - L^* , a^* , b^* » ou plus souvent abrégé « échelle CIELAB ». Ce système CIELAB a les propriétés d'un espace euclidien. On peut donc exprimer chaque point par des coordonnées rectangulaires ou cylindriques (figures 1 et 2).

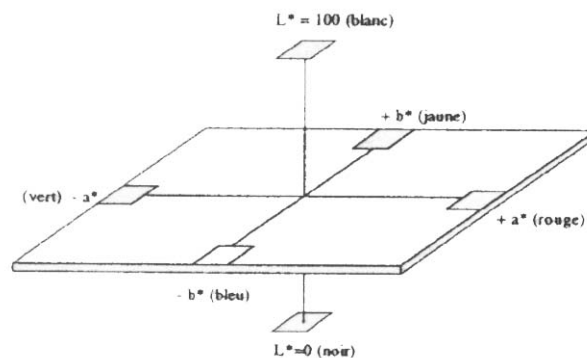


Figure 1 - Représentation CIELAB en coordonnées rectangulaires

CIELAB representation in rectangular coordinates

Le rapport b/a (jaune-bleu/rouge-vert) peut être rapproché de la teinte (rapport jaune/rouge).

On détermine, également :

- la luminosité Y % qui varie de 0 (noir) à 100 p. cent (blanc) ;

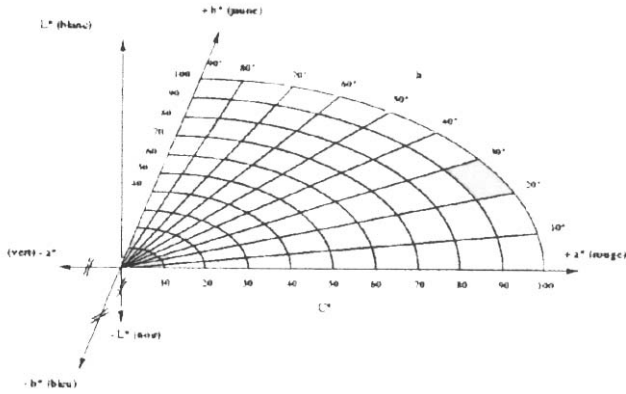


Figure 2 - Représentation CIELAB en coordonnées cylindriques

CIELAB representation in cylindric coordinates

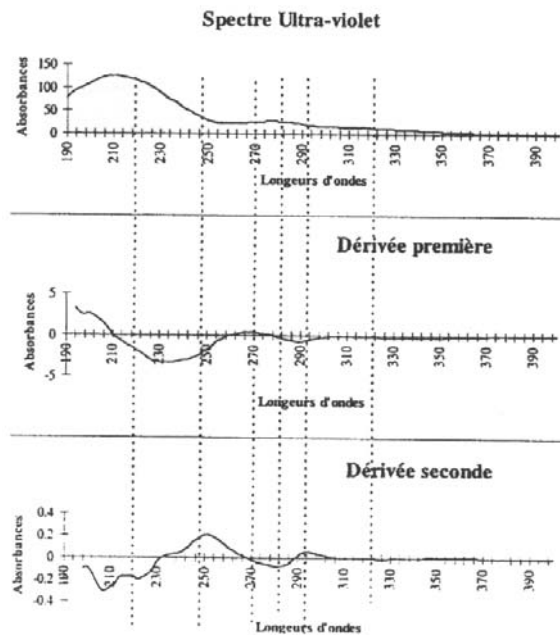


Figure 3 - Absorption dans l'ultraviolet
Absorption in ultraviolet

- la chromaticité par la longueur d'onde dominante LOD, allant de 550 à 680 nm et de 495 à 530 nm pour la LOD dite « complémentaire ».

- la pureté P % qui varie de 0 à 100 p. cent.

En coordonnées cylindriques, on a :

- la clarté L* (identique à L* en coordonnées rectangulaires) ;

- la saturation (ou chroma) C*, définie à partir de a* et b*. $C = (a^{*2} + b^{*2})^{1/2}$ varie de 0 à 141.

- l'angle de teinte (ou tonalité chromatique) h, défini à partir de a* et b*. $h = \arctg(b^*/a^*)$, varie de - 45 à + 45 degrés.

Par ailleurs ces valeurs chiffrées peuvent être exprimées par des mots plus expressifs, tels que rose pourpre, rouge « cerise », « pivoine », « rose hortensia », « rose hibiscus », « rose palmyre » selon les terminologies utilisées dans l'industrie des colorants. Cette technique est exploitée en œnologie (INIGUEZ *et al.*, 1998) pour les vins rouges (rouge violacé, pourpre, grenat, cerise, rubis, tuilé, châtaigne et marron) et rosés (rose framboise, fraise, groseille, saumon et pelure d'oignon) de Rioja.

3) Exploitation directe de l'allure de spectres ultra-violets

Outre la mesure classique d'absorbance à 280 nm qui exprime l'indice des polyphénols totaux (IPT), nous avons calculé la dérivée première et la dérivée seconde des spectres chaque vin afin de mettre en évidence les extrema et les points d'inflexion (figure 3).

Il ressort de cette étude que seuls les facteurs suivants doivent être pris en compte :

Spectre visible	Spectre ultraviolet
ICM	DO 280
Teinte	DO 212 (sommet)
LOD	IPT
Y%	DO 318 (point d'inflexion)
a*	
b*	

TABEAU I
Influence du SO₂, du pH et de l'éthanal sur l'ICM et l'IPT (moyenne de 4 vins)
Influence of SO₂ (mg/l), pH and ethanal on ICM et IPT (average of 4 wines)

	SO ₂ (mg/L)			pH			Ethanal ajouté (mg/L)			
	12	25	45	2,8	3,2	3,7	5	25	50	100
ICM	7,4	6,9	6,4	9,5	8,4	7,6	5,8	7,1	7,4	7,7
IPT	44,4	44,6	44,7	44,9	44,8	44,6	44,4	44,5	44,4	44

Une étude spécifique a montré que ces mesures spectrales sont très répétables (coefficients de variation de 0.8 à 5 p. cent, moyenne 3 p. cent entre 270 et 620 nm) ; il en est de même pour les coordonnées CIELAB (coefficient de variation moyenne = 2.2 p. cent).

Par ailleurs si l'ICM est très sensible au pH ainsi qu'à la teneur en SO₂ ajusté par sulfitage ou addition d'aldéhyde, il apparaît que les valeurs spectrales dans l'ultraviolet sont très stables (tableau I). Ces valeurs ont été complétées par un indice de flavonoïdes totaux (IFT) calculé à partir des absorbances à 280 et 320 nm et correspondant à l'absorbance brute à 280 nm corrigée de l'absorbance due aux protéines, mesurée à 280 nm, après élimination des polyphénols par fixation sur la PVPP (BLOUIN, 1992) :

$$IFT = (A_{280} - 4) - 2/3 (A_{320} - 1.4)$$

4) Approche statistique

Cette technique, reposant sur une décomposition originale des spectres en vecteurs de scalaires, (RICHARD *et al.*, 1999 ; DEHEUVELS *et al.*, 1999 ; Programme ABC, 1994-97) permet d'exploiter la totalité d'un spectre. Par exemple, dans un chromatogramme, on exploite d'une part les pics caractéristiques assez facilement mesurables individuellement et d'autre part les zones bruitées, très caractéristiques, mais dont les pics qui les composent ne sont pas exploitables individuellement, car peu ou pas identifiés et/ou identifiables, non quantifiables.

Les vecteurs représentant les spectres d'absorption lumineuse sont plus simples, mais on n'exploite habituellement que quatre longueurs d'onde (280, 420, 520 et 620 nm), soit une perte d'information considérable correspondant aux très nombreux composés phénoliques (et autres) absorbant sur une large plage de longueurs d'onde.

Un des intérêts de cette approche statistique innovante tient au fait que le modèle représentant les spectres est ajusté en fonction de la base initiale et non préétabli.

En outre, il est possible de compléter ce vecteur en y ajoutant toutes données, y compris des données organoleptiques, ce qui améliore considérablement l'identification et classement.

EXPÉRIMENTATION

Les 213 vins ont été traités d'une part, selon le programme d'essais décrit ci-dessus, et d'autre part, après décomposition vectorielle. Dans ce dernier essai, les pics inférieurs à 260 nm ont été ignorés pour des raisons techniques.

I - COMPARAISON DES DIVERS INDICES CHROMATIQUES

L'étude des tableaux de corrélations entre les indices chromatiques des 213 vins fait apparaître de nombreuses corrélations très hautement significatives.

- La clarté L* est la même pour les deux méthodes CIELAB. Elle est très corrélée à A 420 (r = - 0.92), à ICM (r = -0.91) et à P % (pureté, r = - 0.96) = 0.95) ainsi qu' à Y % (luminosité, r = 0.95).

- A420, A520, A620, IC et ICM sont très corrélées (r > 0.94), malgré leurs différences apparentes. La très haute corrélation entre IC et ICM (r > 0.99 sur ces 213 vins) montre que le rôle réel de la « couleur bleue » est souvent très limité.

- Le nombre de couleur (NbC) est très corrélé à A620 (couleur bleue). Son calcul est basé sur les absorbances à 625, 650 et 675 nm, correspondant à ces couleurs bleues.

- Ces valeurs sont par contre très peu corrélées à la teinte T (r < 0.28) qui est corrélée de façon négative très forte (r = - 0.97) à DA%. Cette situation logique puisque T croît avec la teinte « jaune », alors que DA% exprime le pourcentage de « rouge » dans l'ICM. Notons que le rapport b/a (jaune-rouge/rouge-vert) n'est que peu corrélé avec ces dernières valeurs.

- Très logiquement, il y a une forte corrélation entre h (tonalité chromatique) et b/a (jaune-rouge/rouge-vert) qui sert à son calcul.

- C (saturation) est très corrélé à a (composante vert-rouge) qui intervient dans son calcul.

- IPT et IFT (flavonoïdes totaux estimés par colorimétrie) sont très corrélés (r = 0.99) et semblent mesurer des constituants très proches, probablement parce l'absorbance des protéines des vins rouges, à 280 nm, est faible.

Ces quelques considérations montrent qu'il est possible de simplifier certains bulletins d'analyses en ne retenant que les « indices de couleur » les plus significatifs car plus simples, ou plus parlants.

- L'intensité colorante (IC ou ICM) exprime bien et simplement le niveau de couleur également caractérisable par L (clarté), P (pureté) ou Y (luminosité), plus complexes.

- La teinte T exprime bien la note « tuilée », en opposition logique avec DA p. cent correspondant à la note rouge. Les paramètres b/a, ou h, précisent cette nuance car ils ne sont que très peu corrélés à d'autres paramètres.

- Les composés phénoliques totaux sont bien représentés par l'IPT, plus simple que l'IFT.

- Les paramètres c2 et c3 découlant du traitement complet du spectre ne peuvent pas être reliés à une caractéristique chromatique ou phénolique simple mais ils sont très discriminants, car très faiblement corrélés à tous les autres paramètres étudiés ici (19 et 25 coefficients de corrélation sur 25 inférieurs à 0,20).

II - NORMALISATION DES DONNÉES

1) Traitement préliminaire

A partir des données brutes (figure 4) constituées par deux lots de 213 fichiers correspondant respectivement aux mesures UV et visibles effectuées avec des dilutions différentes, nous reconstituons un fichier unique par vin.

L'ensemble de ces fichiers est normalisé de la façon suivante :

- translation verticale de sorte que chaque fichier ait au moins une ordonnée nulle et toutes les autres positives (ou nulles) ;

- division de chaque ordonnée par la surface totale du spectre, après la translation ;

- lissage du spectre pour éliminer les (éventuelles) perturbations de données.

On représente les spectres après cette normalisation (figure 5).

2) Les empreintes

On transpose la méthode développée sur les chromatogrammes (RICHARD *et al.*, 1999) aux spectres UV-visibles. Le « pic majeur » (environ 280 nm) est éliminé. On calcule par ailleurs un spectre moyen calculé à partir de l'ensemble des fichiers de même nature, puis on calcule un « résidu » :

Résidu = spectre total – pic(s) majeur(s) – spectre moyen sans pic(s) majeur(s)

Empreinte = spectre total – pic(s) majeur(s)

On peut dire que les données de base sont transformées en une « empreinte », par analogie avec les empreintes digitales réduisant les innombrables sillons et autres méandres de nos doigts en quelques données caractéristiques permettant une identification rapide, précise et univoque.

Le résidu est, ici, transformé en un vecteur de 4 composantes calculées de telle sorte que la perte d'information résultant de cette vectorisation ne dépasse pas

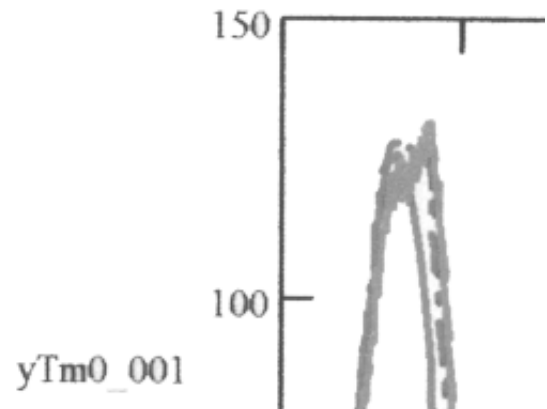


Figure 4 - Représentation de quelques spectres bruts
Representation of some raw spectra

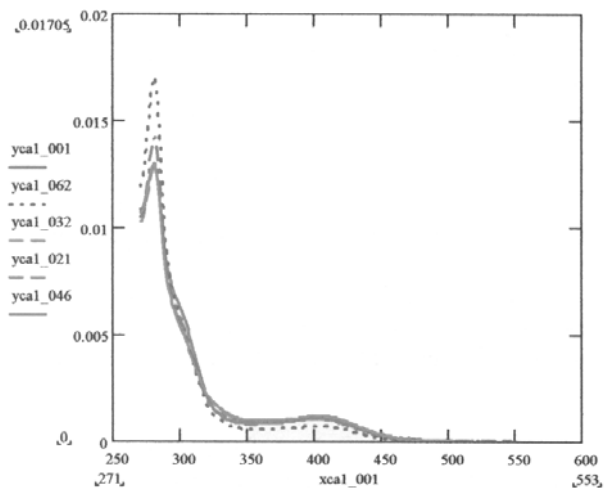


Figure 5 - Spectres bruts après normalisation
Raw spectra after standardization

une valeur α fixée arbitrairement à l'avance. On prend classiquement $\alpha = 5$ p. cent. Cette valeur peut évidemment être différente et si on prenait par exemple $\alpha = 1$ p. cent, on augmenterait le nombre de composantes du vecteur.

Avec $\alpha = 5$ p. cent, chaque spectre peut désormais être représenté par un vecteur à cinq composantes : surface relative du pic majeur et les quatre composantes représentant son empreinte. Quelques empreintes sont représentées figure 6.

L'ensemble des spectres UV-visible complets des 213 vins se trouve réduit à une matrice de 213 lignes et 5 colonnes seulement (tableau I).

Ces quatre composantes permettent la reconstitution de l'empreinte originale avec la perte d'information maximum admise. La figure 7 montre deux exemples de cette reconstitution et les valeurs très

TABLEAU II
Exemples des cinq composantes des dix premiers vins
Examples of the 5 scalar components obtained from the first 10 wines

Numéro du vin	Pic majeur	Composante 1	Composante 2	Composante 3	Composante 4
1	0.79704E+00	0.48739E-03	0.15502E-02	0.77609E-03	-0.21848E-03
2	0.72037E+00	0.54803E-02	-0.85779E-03	0.15501E-02	0.51024E-03
3	0.86654E+00	-0.30102E-02	0.43927E-03	0.13673E-04	-0.11208E-02
4	0.82303E+00	0.40747E-03	-0.17613E-02	0.40342E-03	-0.60348E-03
5	0.82937E+00	-0.31769E-03	-0.27602E-03	-0.12344E-02	-0.33913E-03
6	0.81295E+00	0.10945E-02	-0.16740E-02	0.62069E-03	-0.41702E-03
7	0.76980E+00	0.31978E-02	-0.11613E-02	0.10202E-02	0.23107E-03
8	0.85998E+00	-0.29371E-02	0.37723E-03	-0.31318E-04	-0.11280E-02
9	0.81593E+00	0.93340E-03	-0.17100E-02	0.55258E-03	-0.51934E-03
10	0.81065E+00	0.11848E-02	-0.16537E-02	0.62969E-03	-0.40638E-03

faibles des écarts entre spectre initial et spectre reconstitué.

III - COMPARAISON AVEC LES AUTRES RÉSULTATS

La stabilité des absorbances dans l'ultraviolet montrent les intéressantes possibilités de ces mesures pour caractériser les vins de façon stable, contrairement aux mesures en lumière visible très sensibles aux variations de pH et de teneur en SO₂ libre, par addition ou blocage par les aldéhydes pouvant apparaître lors d'oxydations.

Ces données issues des spectres ultraviolets semblent particulièrement intéressantes comme traceurs d'un vin au cours de sa vie, sur une durée à préciser, mais assez longue.

Si on calcule les corrélations entre les cinq composantes du vecteur (tableau II), sur les 213 vins, on

constate la très forte corrélation négative entre le pic majeur et la première composante ($R^2 = 0,9573$), alors que les autres relations sont très faibles ou pratiquement nulles: la décomposition des spectres en vecteurs met en évidence des caractéristiques n'apparaissant pas lors de l'analyse directe du spectre, même par examen des dérivées successives.

Le tableau de corrélation de tous les paramètres analytiques et des cinq composantes du spectre précise les relations possibles entre ces valeurs.

La première composante est très corrélée à A520, IC, ICM et le pic majeur: elle correspond à l'intensité colorante et aux polyphénols totaux.

Le fait que la corrélation entre le « pic majeur » et la première composante de l'empreinte soit très proche de 1 suggère qu'on réduise à 4 le nombre de composantes du vecteur final ce qui signifierait que la partie empreinte seule est suffisante pour expliquer tout le spectre.

On vérifie que chacune des 4 composantes de l'empreinte est bien indépendante des 3 autres, conformément à la méthode même de calcul.

Les autres composantes du vecteur empreinte sont pratiquement indépendantes des valeurs obtenues directement à partir du spectre.

On peut reconstituer chaque empreinte uniquement à partir des composantes c_1, c_2, c_3, c_4 de son vecteur. Nous avons représenté figure 7 la reconstruction de deux empreintes. La courbe en rouge, appelée yca7a_001, est l'empreinte brute, la courbe en pointillé bleu, appelée yca7b_001, est l'empreinte reconstituée, la courbe en vert, appelée yca7c_001, représente la différence, point par point, entre l'empreinte brute et l'empreinte reconstituée.

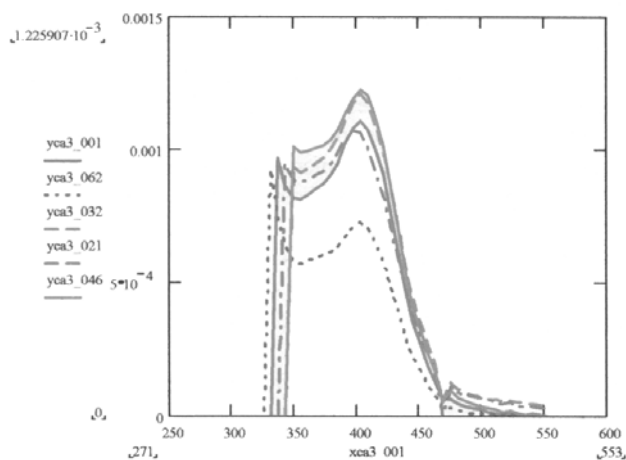


Figure 6 - Quelques empreintes.
On notera le changement d'échelle des ordonnées
Some fingerprints. Note the important change of scale

On a accepté une perte d'information de 5 p. cent maximum. La reconstitution de droite (yca7b_114) est moins bonne que celle de gauche (yca7b_001). On pourrait donc se demander si il n'aurait pas été judicieux d'éliminer le n°114 [].

IV - CLASSIFICATIONS

La richesse des données fournie par un spectre UV-visible de vins rouges incite à comparer ces données.

La méthode utilisée sera détaillée dans une autre publication, elle se distingue des méthodes classiques de classifications multivariées et peut être ainsi résumée :

Les individus sont aléatoirement répartis en C classes arbitrairement fixées. On choisit C de telle sorte qu'il soit nettement supérieur au nombre probable de classes. Pour chaque individu, on calcule ensuite les écarts entre cet individu et chacun des autres groupes, pour tous les paramètres (ici les 5 composantes du vecteur). On affecte ensuite cet individu au groupe le plus proche et on refait le calcul. L'expérience montre une rapide convergence des calculs avec définition de groupes stables et homogènes, établis sans aucun *a priori*, pour tous les caractères à la fois. Le nombre final de classes est automatiquement défini et inférieur à C, si C a été choisi suffisamment grand.

Des tests spécifiques, issus des travaux de Wilks (KRES, 1983), et des tables correspondantes complétées par nos soins, permettent la mise en évidence de données et de résultats aberrants, au sens statistique, c'est-à-dire s'écartant fortement de la population générale. Après examen individuel, on peut ainsi détecter les erreurs de transcriptions de données, les mesures

erronées, les individus effectivement aberrants (exemple : un vin blanc dans un groupe de vins rouges...) ou aux propriétés exceptionnelles (exemple : dégustateur à la sensibilité exacerbée pour tel stimulus...).

Les 213 vins ont ainsi été classés en 14 groupes, constitués de 6 à 40 individus. Quelques exemples :

- Le groupe 7 est exclusivement constitué de 11 vins (Médoc, 1997) différents seulement par leur teneur en SO₂ libre ou des additions d'aldéhydes, mais les deux vins aux pH modifiés sont séparés, dans le groupe 8.
- Il en est de même pour le groupe 6 pour un vin (Merlot, 1997) traité de la même façon, mais les deux vins très sulfités sont isolés dans le groupe 5.
- Les groupes 13 et 14 sont chacun constitués de 1 ou 2 familles de vins de même origine.
- *A contrario*, un groupe ayant subi les mêmes modifications est réparti entre les groupes 2 et 4, ce dernier regroupant les vins sulfités.
- Tous les rosés et clarets sont regroupés en deux groupes distincts (9 et 10) alors que l'examen des critères usuels (IPT, ICM, T) ne distingue qu'un seul groupe hétérogène.
- 8 des 10 vins du groupe 1 sont de même origine (Merlot noir, 1998) et plus ou moins tanisés. Ils sont accompagnés de 2 vins « vieux » (1994 et 1991).
- Les groupes 3, 4 et 8 sont plus diversifiés, constitués de sous-groupes variés.

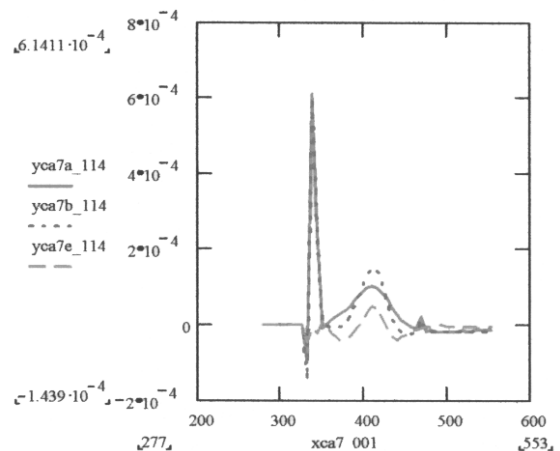
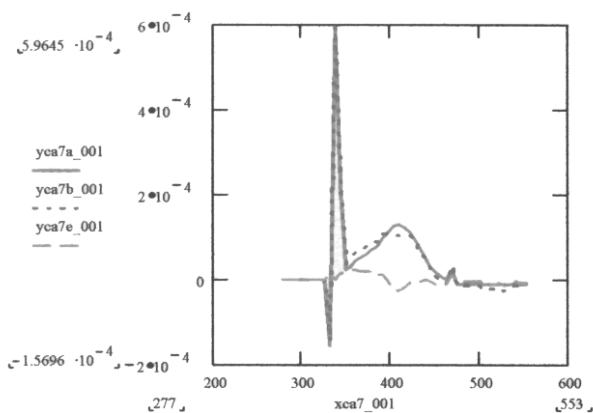


Figure 7 - Deux empreintes reconstituées.

On remarquera que l'empreinte de droite (114) pourrait sûrement être éliminée

Two reconstituted fingerprints.

Note that the right fingerprint (114) could be to be eliminated

DISCUSSION

Quelles peuvent être les causes des différences ?

Pourquoi tel vin d'une famille supposée homogène (exemple : même origine) est-il différent des autres ?

Cette technique peut être d'application très rapide (un spectre complet est obtenu en quelques secondes), à coût aujourd'hui modéré, pour des classifications de travail.

En détectant les situations « aberrantes », la classification automatique attire l'attention sur des données méritant des explications complémentaires, avec ou sans analyses complémentaires. L'analyse du spectre infrarouge est également possible et prend en compte de très nombreux autres paramètres analytiques complétant les possibilités de classification. Notons enfin que les vecteurs issus de ces spectres peuvent être complétés par des données analytiques traditionnelles et des données issues de la dégustation (exemple : notes pour différents descripteurs).

Le vin A est-il (ou non) identique, très proche ou très éloigné du vin B, au seuil de confiance (95 ou 99 p. cent par exemple choisi ?

La tentation est parfois grande de considérer successivement chaque paramètre décrivant les vins et de comparer une à une les différentes valeurs mesurées ou calculées et de les admettre comme identiques (ou différentes) avec une probabilité X p. cent, par exemple de 95 p. cent (ou 99 p. cent...). La réflexion et la théorie mathématique (RICHARD et DEHEUELS, 1998) montrent qu'avec l'incertitude X pour chaque paramètre, l'incertitude totale est comprise entre $1-X$ et $P(1-X)$ pour P paramètres. Ainsi avec dix paramètres et une probabilité de 95 p. cent, l'incertitude serait comprise entre 5 et 50 p. cent. Il ne faut donc pas confondre une procédure multicritère, où plusieurs critères sont examinés successivement, avec une procédure multivariée exploitant globalement l'ensemble des paramètres, avec une probabilité globale fixée à l'avance, 95 p. cent par exemple. A titre d'exemple, cette méthodologie a été appliquée pour distinguer des jus de fruits selon leurs traitements thermiques (Programma ABC, 1994-1997).

CONCLUSION

Le regroupement automatique et impersonnel de vins diversifiés à partir de la seule mesure de leur spectre UV-visible, ouvre de très importantes perspectives pratiques.

Le vin A est-il - ou non - identique, au seuil de confiance choisi, au vin B ?

Dans un vaste ensemble de vins, on peut ainsi mettre en évidence des parentés de structure, de couleur et/ou de composés phénoliques.

Les mesures obtenues par une autre voie (exploitation directe des valeurs spectrales ou mesures organoleptiques) peuvent être ajoutées au vecteur empreinte et permettre une exploitation multivariée globale.

Cette technique originale de traitement du signal permet en outre de réduire considérablement le volume de stockage de données d'origines chromatographiques, spectrophotométriques UV-visible, comme ici, ou infrarouge. Elles peuvent être exploitées, reconstituées pour tous usages, dont la classification.

RÉFÉRENCES BIBLIOGRAPHIQUES

- AFNOR., 1984. Couleur, colorimétrie. *Recueil des normes françaises NF X08-013*, p 514-520.
- BLOUIN J., 1992. *Techniques d'analyses des moûts et des vins*. Laboratoires Dujardin-Salleron éd., Paris.
- DEHEUELS P. PIERRE-LOTI-VIAUD D et RICHARD J.P., 1999. *A new approach to the interpretation of chromatograms. Technical Reports*. ABC, Paris.
- DUJARDIN-SALLERON., 1928. *Notice sur les instruments de précision appliqués à l'œnologie*. 6^e édition, p 629.
- FAURE A et PALLU R., 1935. *Ann. Fals. Fraudes*, **28**, 5.
- FLANZY M et POUX C., 1958. *Ann. Technol. Agric.*, **7**, 377.
- GLORIES Y., 1984. La couleur des vins rouges. *Connaissance Vigne Vin*, **18**, 3, 195.
- INIGUEZ CRESPO M., 1998. Control de calidad en el color de uvas y vinos. Sencillez zn los métodos anlisis. *Jornads Tecnicas*, Universidad Rovira i Virgili, Lallemand, Tarragona, 39-44.
- JOURNAL OFFICIEL CEE., 1990. Analyses des moûts et des vins. *Règlement CEE n° 2676/90*.
- KRES H., 1983. *The multivariate outlier criteria of S.S WILKS*. Springer Verlag ed.
- PROGRAMME AIR CT-94-1111 – 1994-1997 – DG XII, 8, square de Meeus 1027 Bruxelles.
- RIBÉREAU-GAYON P., 1970. *Chim. Anal.*, **52**, 627.
- RICHARD J.P., DEHEUELS P. et PIERRE-LOTI-VIAUD D., 1999. Une nouvelle approche pour l'exploitation des empreintes chromatographiques. *Ann. Fals. Exp. Chim.*, **92**, 455.
- RICHARD J.P. et DEHEUELS P., 1998. *Estimation de la fidélité et des régions de confiance pour les données multivariées*. Programme AIR, 1998.
- ROOS., 1930. *Ann. Fals. Fraudes*, **23**, 207.

Reçu le 30 août 2001

accepté après révision le 24 janvier 2002

ROUBERT J., 1977. Détermination de la couleur par la

méthode des tristimuli. *La France Viticole*, **3**,

71-72.

SUDRAUD P., 1958. *Ann. Technol. Agric.*, **7**, 203.