

Multi-mineral wine profiling and Artificial Intelligence: Implementing the signatures of each wine to train algorithms to meet the new challenges facing the wine industry

Augustin Tillement^{1,2,3}, Coraline Duroux², Leticia Sarlo^{1,2}, Théodore Tillement², Laurent David³, and Olivier Tillement^{1,2}

¹ Institut Lumière-Matière, UMR 5306, Université Claude Bernard Lyon 1-CNRS, Université de Lyon, Villeurbanne Cedex 69100, France

² M&Wine, 305 rue des Fours, 69270 Fontaines Saint Martin, France

³ Institut National des Sciences Appliquées, CNRS UMR 5223, Ingénierie des Matériaux Polymères, Université Claude Bernard Lyon 1, Université Jean Monnet, 15 bd Latarjet, 69622 Villeurbanne, France

Abstract. Although present in minimal quantities, minerals are essential to the composition of every wine. Their presence results from complex interactions between various factors such as soil, vines, climate, topography, and viticultural practices, all strongly influenced by terroir. Each stage of the winemaking process further shapes the unique mineral and flavor profile of each wine. This mineral composition, stable over time, can serve as a foundation for developing new artificial intelligence (AI) algorithms. To this end, we have established a comprehensive database of over 20,000 wines, each characterized by a "Mineral Wine Profile" (MWP), a concentration profile of forty mineral elements obtained through inductively coupled plasma mass spectrometry (ICP-MS). These profiles constitute robust and extensive datasets, ideal for integration into AI-guided oenology. Beyond the specific signature of a wine and the development of authentication services, combining these signatures with additional data allows for the creation of new AI-driven tools to address emerging challenges in the wine industry. We will present the first proofs of concept that demonstrate the potential of this AI and mineral analysis approach for the creation of targeted wines.

1. Introduction

La prédiction de la perception sensorielle et de l'appréciation des caractéristiques organoleptiques du vin par les consommateurs représente un défi majeur en œnologie. Développer des modèles prédictifs précis pour anticiper les attributs sensoriels des vins pourrait transformer la production viticole en permettant une personnalisation accrue en fonction des préférences des consommateurs [1,2]. Outre cette adaptation aux nouveaux modes de consommation, ces modèles pourraient offrir de nombreuses applications, telles que l'amélioration du contrôle qualité, l'authentification des produits [3,4,5,6], l'adaptation aux changements climatiques et viticoles ainsi que l'innovation dans le développement de nouveaux vins [7].

En tant que donnée massive pour l'intelligence artificielle (IA), le profil minéral du vin (MWP) constitue une empreinte stable et distinctive qui a déjà démontré son potentiel pour l'authentification des vins [9,10]. Dans cette

étude, nous avons employé la spectrométrie de masse à plasma à couplage inductif (ICP-MS) pour mesurer de manière relative la concentration des éléments suivants : ¹¹B, ²³Na, ²⁴Mg, ²⁷Al, ²⁸Si, ³¹P, ³⁴S, ³⁵Cl, ³⁹K, ⁴³Ca, ⁴⁵Sc, ⁴⁷Ti, ⁵¹V, ⁵²Cr, ⁵⁵Mn, ⁵⁶Fe, ⁵⁹Co, ⁶⁰Ni, ⁶³Cu, ⁶⁶Zn, ⁷⁵As, ⁷⁹Br, ⁸⁵Rb, ⁸⁸Sr, ⁸⁹Y, ⁹⁰Zr, ⁹³Nb, ¹¹¹Cd, ¹¹⁸Sn, ¹²⁷I, ¹³³Cs, ¹³⁷Ba, ¹³⁹La, ¹⁴⁰Ce, ¹⁴¹Pr, ¹⁴⁶Nd, ¹⁸²W, ²⁰⁵Tl, ²⁰⁸Pb, ²³⁸U [3,11]. Ce profil minéral est influencé par la composition du sol, les techniques de vinification, les cépages, et divers autres paramètres allant de la vigne à la bouteille [8,12,13]. Alors que les profils minéraux ont montré des résultats prometteurs pour l'identification et l'authentification des vins [10,14], notre étude explore leur potentiel innovant pour prédire les caractéristiques organoleptiques. En associant ces profils minéraux aux descriptions sensorielles, nous ouvrons la voie à une meilleure compréhension et prédiction des attributs sensoriels complexes du vin [15].

Les attributs sensoriels des vins sont souvent perçus de manière émotionnelle et subjective [16]. Des panels de consommateurs ou d'experts peuvent évaluer ou comparer

des vins, tandis que des données additionnelles peuvent être recueillies à partir de différentes bases de données publiées [17]. Bien que ces informations puissent être biaisées par les préférences des consommateurs, influencées par des facteurs tels que le prix ou les avis antérieurs [18], ces bases de données offrent néanmoins des avantages substantiels. L'accumulation d'un volume important de données renforce la puissance statistique, permettant d'identifier les qualités, profils sensoriels et principales caractéristiques organoleptiques des vins [19]. Nos recherches reposent sur l'hypothèse que la combinaison du profil minéral et sensoriel d'un vin permettrait de développer des modèles prédictifs robustes, capables de fournir des informations essentielles pour l'industrie viticole [20].

De nombreuses études ont tenté de prédire les caractéristiques organoleptiques des produits alimentaires à partir de leurs propriétés chimiques, mais cela reste un défi significatif en science sensorielle, en particulier pour les aliments et boissons complexes tels que le vin [13, 21]. Un obstacle majeur réside dans la diversité et la complexité des composés aromatiques actifs, dont la structure moléculaire et la concentration peuvent varier de manière significative [5,21]. Cette variabilité rend leur quantification techniquement ardue, même face aux avancées en métabolomique, telles que le fingerprinting métabolique non ciblé [23]. De plus, l'analyse sensorielle est d'autant plus complexe que la perception des saveurs résulte de l'interaction de centaines de molécules à la fois au niveau physico-chimique et sensoriel [24]. Cette perception est souvent non linéaire, marquée par des effets synergiques et antagonistes en fonction de la concentration, et encore compliquée par des facteurs génétiques, environnementaux, culturels et psychologiques des consommateurs [25]. Jusqu'à présent, la prédiction des propriétés organoleptiques s'est principalement concentrée sur des composés isolés [15].

L'étude du profil minéral, bien que moins directement liée aux goûts et aux arômes que d'autres composés chimiques, présente trois avantages significatifs : un accès simple aux analyses couplé à une stabilité des éléments au fil du temps [26] et des associations complexes indirectes avec diverses molécules et processus viticoles et vinicoles [27], ainsi que leur résilience aux conditions de stockage et de vieillissement [28]. Ces profils minéraux constituent des données robustes et massives (*Massive Data*) porteuses d'informations complexes. L'utilisation de méthodes statistiques multivariées et de techniques de *machine learning* sur ces profils minéraux pourrait constituer une approche fiable pour prédire certains attributs sensoriels du vin [29].

Dans cette étude, nous exploitons les profils minéraux pour mener une analyse globale des attributs sensoriels du vin et développer des modèles prédictifs. Notre base de données, comprenant plus de 20.000 MWP d'échantillons de vins provenant de plus de 200 régions viticoles et couvrant plus de 150 cépages [19], offre une diversité exceptionnelle pour l'analyse. Parmi ces données, près de 8.000 vins incluent des informations qualitatives, aromatiques et/ou gustatives exploitables [30]. En

combinant cette richesse de données avec des techniques avancées de *machine learning*, nous développons des modèles capables de prédire avec précision une large gamme d'attributs organoleptiques, tels que l'intensité, l'acidité, la douceur (ou sucrosité), ainsi que des notes spécifiques comme les arômes végétaux, floraux, fruités, épicés et de minéralité [7].

L'intégration de ces données avec des descriptions sensorielles permet de relier la composition minérale aux attributs organoleptiques, ouvrant ainsi la voie à de nouvelles applications dans l'industrie viticole [31]. Cette approche novatrice pourrait accompagner l'évaluation sensorielle et faciliter le développement d'outils prédictifs pour le secteur vitivinicole [32].

2. Matériel et méthode

2.1. Réactifs et matériaux

Au total, 20.000 vins issus de compétitions internationales, ainsi que de vins commerciaux, provenant de plus de 200 régions viticoles et couvrant plus de 150 cépages différents ont été analysés dans cette étude. Pour chaque vin, environ 30 mL ont été prélevés dans des tubes certifiés sans métal (VWR®). Les échantillons ont été dilués dans un rapport de 1:3 avec de l'acide nitrique à 1 % (v/v), préparé avec de l'eau ultrapure (MilliQ®, 18,2 mΩ.cm) et de l'acide nitrique de qualité Suprapur® (69 % v/v, Roth), auquel a été ajouté une solution standard d'indium à 10 µg/L, préparée avec une solution d'indium standard à 1000 mg/L dans HNO₃ à 4 % (v/v). Cette dilution permet de conserver les échantillons de vin dans des conditions acides, préservant ainsi le profil minéral du vin (MWP) au fil du temps en empêchant la précipitation et l'adsorption sur les parois des tubes. Avant l'analyse, une seconde dilution de 1:5 a été réalisée avec HNO₃ à 1 % (v/v), en utilisant de l'acide nitrique de qualité Suprapur® (69 % v/v, Roth) dilué dans de l'eau ultrapure (MilliQ®, 18,2 mΩ.cm). Ce facteur de dilution (1:15) a été démontré comme étant efficace pour minimiser les effets de matrice [10,32].

2.2. Détermination du profil minéral du vin (MWP)

Les profils minéraux des vins (MWP) ont été déterminés par des analyses ICP-MS, réalisées entre juin 2022 et juillet 2024 chez M&Wine, Lyon-France, et à l'Institut des Sciences Analytiques, Université Claude Bernard Lyon 1, en utilisant différents équipements ICP-MS quadripôles. La majorité des mesures a été effectuée avec l'ICP-MS quadripôle simple *Agilent Technologies 7850*, équipé d'un auto-échantillonneur intégré SPS 4. Un nébuliseur micromist a été utilisé pour toutes les mesures. Pour minimiser les interférences polyatomiques, la cellule de collision a été réglée en mode Hélium pour tous les éléments, avec un débit de 5 mL/min. Les paramètres de fonctionnement étaient les suivants : puissance avant de 1550 W, débit du gaz porteur de 1 L/min, débit du gaz plasma de 15 L/min et débit du gaz auxiliaire de 1 L/min. Une solution d'ajustement contenant 1 µg/L de Ce, Co, Li,

Tl, et Y dans HNO₃ à 2 % (v/v) (*Agilent Technologies*) a été utilisée en début d'analyse. Un étalon multi-éléments (VWR®, référence 85006.186) contenant 100 mg/L de Ag, As, B, Ba, Be, Bi, Ca, Cd, Co, Cr, Cu, Fe, K, Li, Mg, Mn, Mo, Na, Ni, Pb, Sb, Se, Sr, Ti, Tl, V, et Zn dans HNO₃ à 5 % (v/v) a été dilué pour obtenir l'étalon de calibration semi-quantitative à une concentration de 20 µg/L. Pour contrôler et valider cette calibration, un vin d'origine commerciale a été utilisé comme référence, préparé selon la même procédure que celle utilisée pour les échantillons de vin décrits précédemment. L'analyse du vin de concentration connue a été réalisée au début, au milieu et à la fin de chaque séquence d'échantillons. Les blancs et l'étalon à 28 éléments ont été réanalysés tous les 40 échantillons. L'analyse semi-quantitative (SQ) a permis de déterminer les concentrations élémentaires en utilisant l'étalon à 28 éléments. Les 40 éléments précédents ont été analysés, avec 100 balayages et une réplique ; tous, à l'exception de ¹¹⁵In, l'étalon interne, constituent le MWP.

2.3. Collecte et traitement des données organoleptiques

Des données organoleptiques ont été recueillies sur plus de 8000 vins. Les descriptions sensorielles des vins ont été extraites et regroupées selon les différentes catégories conventionnelles : *Végétale, Florale, Fruits rouges ou noirs, Fruits blancs, Agrumes, Fruits exotiques, Épices et aromates, Boisée, Secs, Balsamique, Emphyreumatique, Animale, Minérale, Lactée, Pâtisserie, Amylique, Défauts*. Au total, plus de 1.650.000 mentions ont été utilisées sur l'ensemble des vins analysés. Les mentions pour chaque catégorie ont été agrégées et normalisées en pourcentage pour éviter qu'un vin très populaire, avec de nombreuses évaluations, ne biaise les informations comparativement à des vins moins connus avec moins de notes. Ce traitement a permis d'obtenir une base de données organoleptique robuste et équilibrée, qui a été utilisée en complément des profils minéraux pour les analyses subséquentes. Nous avons aussi extrait des informations structurelles précises sur les vins autour des quatre axes de goût, à savoir *Puissance (Léger à Puissant), Structure (Souple à Tannique), Texture (Sec à Moelleux)*, et *Douceur (Doux à Acide)*. Une valeur quantitative à ces critères allant de 1 à 5 a été attribuée afin de les intégrer dans nos modèles prédictifs. La *Structure* n'est disponible que pour les vins rouges. Pour les vins effervescents, un autre axe est également accessible, l'*Effervescence* (Tranquille à Mousseux). Nous avons ensuite segmenté ces données en trois catégories distinctes, correspondant respectivement au tiers inférieur, au tiers médian, et au tiers supérieur pour chacune de ces cinq caractéristiques.

2.4. Modèles de Machine Learning

2.4.1. Modèles de classifications

Des modèles de classification par intelligence artificielle (IA) ont été construits pour prédire la présence ou l'absence de profils sensoriels spécifiques dans les vins en

se basant sur leur profil minéral. Les profils organoleptiques ont été classés en deux groupes : les vins sans mention d'un profil sensoriel particulier (Groupe 1) et les vins avec au moins 5 % de mentions pour ce profil (Groupe 2). Les vins pour lesquels le nombre de mentions était strictement positif mais représentait moins de 5 % du total des mentions n'ont pas été pris en compte dans l'analyse. Un pourcentage aussi faible sur un panel non professionnel n'est pas suffisant pour considérer la catégorie sensorielle comme véritablement présente dans le vin, mais reste trop élevé pour être considéré comme nul. Ainsi, ces cas intermédiaires ont été exclus pour garantir la précision et la fiabilité des résultats. Avant la modélisation, des étapes de prétraitement des données ont été effectuées, notamment le filtrage des colonnes comportant plus de 50 % de valeurs en dessous de la limite de détection (LOD). Pour les autres colonnes, avec moins de 50% de valeurs en dessous de la limite de détection, ces valeurs ont été remplacées par des 0. Les concentrations des minéraux ont été normalisées à l'aide d'une mise à l'échelle standard, et les données ont été réparties aléatoirement en ensembles d'entraînement et de test en utilisant une approche stratifiée pour garantir une représentation proportionnelle de chaque groupe de profils sensoriels. Les cinq modèles de *machine learning* suivants ont été mis en œuvre à l'aide de Python (v3.9.16) avec les packages *scikit-learn* (v1.2.2) et *xgboost* (v1.7.3) : *Random Forest* (RF), *Régression Logistique* (LR), *Gradient Boosting Machine* (GBM), *Extra Trees*, *XGBoost* (XGB), *CatBoost*. Chaque modèle a été évalué à l'aide d'une validation croisée stratifiée en cinq volets, où l'ensemble de données a été divisé à plusieurs reprises en ensembles d'entraînement et de validation pour évaluer les performances du modèle sur différents échantillons. L'élimination récursive des caractéristiques (RFE) a été appliquée à chaque modèle pour identifier les caractéristiques les plus pertinentes pour la prédiction. Le nombre de caractéristiques sélectionnées variait entre les modèles afin d'optimiser l'Aire Sous la Courbe (AUC) de la courbe ROC (Receiver Operating Characteristic) [34]. L'AUC est une mesure clé en apprentissage automatique, particulièrement dans les modèles de classification. Elle représente la probabilité qu'un modèle classe correctement un échantillon positif aléatoire au-dessus d'un échantillon négatif aléatoire. Une AUC de 0,5 indique une performance équivalente à celle du hasard, tandis qu'une AUC de 1,0 reflète une classification parfaite. Dans notre étude, nous avons calculé la moyenne des AUC pour chaque modèle à travers toutes les caractéristiques organoleptiques. Le modèle présentant la meilleure moyenne d'AUC a ensuite été sélectionné comme le plus performant. Cette approche nous assure que le modèle choisi est non seulement précis pour une caractéristique spécifique, mais qu'il offre également une robustesse et une capacité de généralisation accrues lorsqu'il est appliqué à l'ensemble des caractéristiques étudiées.

2.4.2. Modèles de Régressions

Nous avons pu appliquer ces techniques de régression uniquement aux profils structuraux, car ce sont les seules caractéristiques pour lesquelles nous disposons de notes sur une échelle de 0 à 5. En revanche, pour les caractéristiques organoleptiques, nous ne disposons que de données binaires indiquant la présence ou l'absence de ces profils, basées sur le taux de mentions correspondant à chaque vin. Étant donné la subjectivité inhérente à ces évaluations, il serait peu pertinent de tenter de prédire des valeurs continues par régression. De plus, dans le contexte vitivinicole, il est souvent plus pertinent et exploitable de prédire la présence ou l'absence de ces caractéristiques organoleptiques plutôt que de fournir un score qui pourrait être interprété de manière ambiguë. Cette approche binaire permet ainsi d'offrir des prédictions claires et plus facilement exploitables par les professionnels du secteur.

Les modèles de régression suivants ont été mis en œuvre en utilisant les mêmes packages que pour les modèles de *machine learning* : **Random Forest Regressor (RFR)**, **Gradient Boosting Regressor (GBR)**, **AdaBoost Regressor (ABR)**, **Régression Linéaire (LR)**, **Lasso Regression**, **Support Vector Regression (SVR)**, **Extra Trees Regressor**, **XGBoost Regressor**, **CatBoost Regressor**. Les performances de ces modèles ont été évaluées en termes de coefficient de détermination (R^2) et de coefficient de corrélation de Spearman (ρ) pour chaque caractéristique organoleptique étudiée. La validation croisée stratifiée a été utilisée pour garantir que les modèles sont robustes et généralisables. Les résultats ont été visualisés à l'aide de courbes de prédiction versus réalité, permettant une interprétation claire de la performance des modèles et des caractéristiques les plus influentes sur les prédictions.

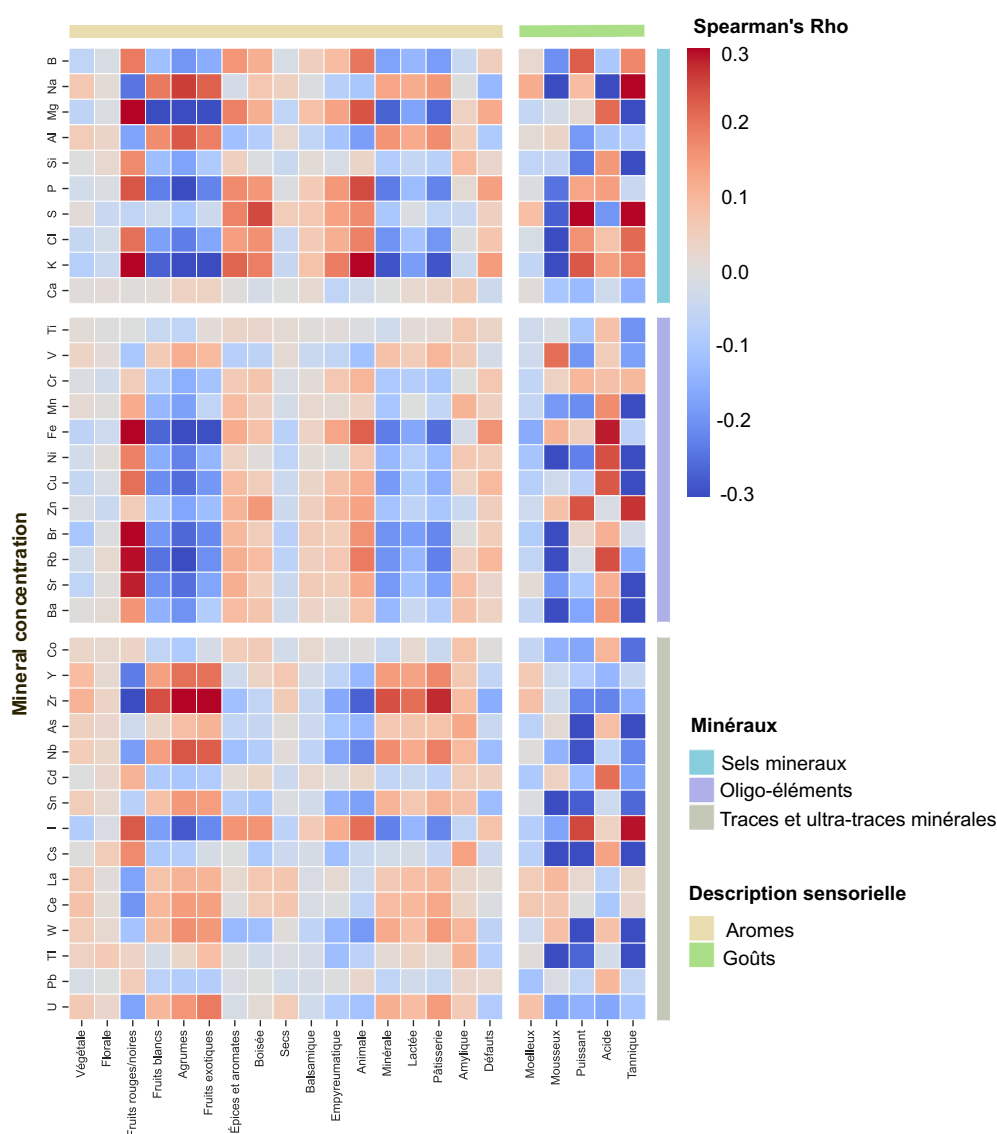


Figure 1. Heatmap de corrélation entre les minéraux et les données organoleptiques La figure présente une heatmap illustrant les corrélations de rang de Spearman entre les concentrations minérales et les descripteurs sensoriels des vins. Les descripteurs sensoriels sont organisés en deux catégories principales : arômes (en vert) et goûts (en beige), permettant de visualiser distinctement comment les différents éléments minéraux influencent ces aspects sensoriels. Les éléments minéraux ont été classés en trois groupes distincts : les sels minéraux (en bleu clair), les oligo-éléments (en violet) et les traces et ultra-traces minérales (en gris). Cette classification permet de mieux comprendre les rôles spécifiques que jouent ces éléments dans la modulation des arômes et des goûts. Les zones de couleur rouge indiquent des corrélations positives, tandis que les zones bleues représentent des corrélations négatives, offrant ainsi un aperçu rapide des relations entre chaque élément minéral et les caractéristiques organoleptiques.

3. Résultats

3.1. Analyse de la composition élémentaire dans les vins

Cette étude de prédiction de qualité s'appuie sur une partie de la base de données comprenant 7.787 profils minéraux de vins (MWP). Les échantillons proviennent de 42 pays différents, 207 régions viticoles, et couvrent 169 cépages distincts. Le Chardonnay et le Merlot représentent les cépages les plus fréquemment analysés, avec respectivement 1.236 et 1.234 échantillons, tandis que Bordeaux et Beaujolais sont les régions viticoles les plus représentées avec 1.663 et 991 échantillons. Quant aux pays, la France domine largement avec 6.531 échantillons, suivie de l'Espagne (187) et de l'Italie (126). Les valeurs observées dans notre ensemble de données concordent avec les plages de concentrations rapportées dans la littérature pour les éléments macro, micro et traces [35]. Cette diversité dans l'origine et les méthodes de production des échantillons renforce la pertinence de notre étude pour le développement de modèles prédictifs robustes des caractéristiques organoleptiques des vins.

3.2. Analyse approfondie des relations entre les composés minéraux et les caractéristiques organoleptiques des vins

Pour chaque vin de cette étude, en plus du profil minéral MWP, nous avons collecté des données organoleptiques, couvrant un large éventail de descripteurs sensoriels. Dans un premier temps, nous avons réalisé une *heatmap* de corrélation (voir Fig. 1). pour identifier les relations entre les concentrations en minéraux et les différentes données organoleptiques des vins. Pour ce faire, nous avons utilisé le coefficient de corrélation de Spearman, qui est particulièrement adapté à l'analyse des relations monotones, même non linéaires, entre les variables. L'analyse a été effectuée sur l'ensemble de la base de données, à l'exception des données spécifiques telles que Moelleux, Mousseux et Tannique. Pour ces catégories, des ajustements ont été réalisés : les descripteurs Moelleux et Mousseux ont été spécifiquement analysés pour les vins effervescents, tandis que Tannique a été examiné uniquement pour les vins rouges. Cette approche permet d'assurer que les relations calculées reflètent des relations pertinentes et spécifiques au type de vin considéré. Les corrélations de rang de Spearman sont présentées (Fig. 1), les descripteurs sensoriels sont regroupés en deux catégories principales : arômes ou perceptions olfactives (vert) et goûts ou perceptions gustatives (beige). Les minéraux sont classés en trois groupes : sels minéraux (bleu clair), oligo-éléments (violet), et traces et ultra-traces minérales (gris). Les données chimiques sont disponibles en annexe, ainsi que les valeurs de corrélation pour toutes les paires de descripteurs sensoriels et minéraux.

3.3. Analyse des caractéristiques structurales

Dans un second temps, l'accent a été mis sur l'analyse des caractéristiques structurales des vins, à savoir la **Puissance** (Léger à Puissant), la **Structure** (Souple à

Tannique), la **Texture** (Sec à Moelleux), et la **Douceur** (Doux à Acide). Ces caractéristiques ont été sélectionnées en raison de leur importance dans l'expérience de dégustation et leur pertinence pour une large gamme de vins. Pour quantifier ces critères, une échelle de 1 à 5 a été utilisée, où 1 caractérise les vins positionnés dans les 20% en tête pour le premier critère. Cela a permis de transformer des perceptions sensorielles subjectives en données quantitatives intégrables dans nos modèles prédictifs.

3.3.1. Sélection de l'algorithme de *Machine Learning*

Afin de sélectionner la stratégie de *machine learning* la plus adaptée, les coefficients de détermination (R^2) et les coefficients de corrélation de Spearman (ρ) ont été calculés pour chaque modèle à travers cinq descripteurs organoleptiques : **Texture (Sec à Moelleux)**, **Effervescence (Tranquille à Mousseux)**, **Puissance (Léger à Puissant)**, **Douceur (Doux à Acide)** et **Structure (Souple à Tannique)**. Les résultats ont été moyennés sur plusieurs essais et sont présentés (Tab. 1). Les résultats indiquent que les profils minéraux constituent d'excellents indicateurs pour prédire ces attributs sensoriels, avec des coefficients de détermination (R^2) et des corrélations de Spearman (ρ) élevés. Le modèle CatBoost Regressor a été sélectionné pour poursuivre les analyses sur ces descripteurs organoleptiques, car il offre en moyenne les meilleures performances.

3.3.2. Analyse des résultats et impact des éléments minéraux

Les graphiques (Fig. 2) offrent une analyse plus visuelle et détaillée des résultats obtenus par le modèle CatBoost Regressor. Le graphique de dispersion permet de comprendre la relation entre les valeurs réelles et les prédictions du modèle pour chaque descripteur (Fig. 2A). Les coefficients R^2 et ρ dans ces graphiques confirment les valeurs du tableau, mais l'intérêt réside dans la visualisation de la répartition des erreurs de prédiction et des tendances capturées par le modèle. Le diagramme SHAP (Fig. 2B) permet d'aller au-delà du simple score de performance en offrant une vue sur les contributions spécifiques des différents minéraux aux prédictions. Cela aide à justifier la fiabilité des prédictions, selon la force de l'influence de ces minéraux. Les résultats pour les différentes caractéristiques étudiées se résument ainsi :

Puissance (Léger à Puissant, vins rouges), $R^2 = 0.68$, $\rho = 0.76$: La prédiction de la puissance, qui décrit l'intensité générale du vin, est particulièrement précise. Le césium (Cs), le nickel (Ni), et le soufre (S) jouent un rôle crucial dans cette perception. Ces éléments, liés aux processus de maturation et d'extraction phénolique, contribuent à renforcer la structure et la complexité des vins rouges. Leur influence notable dans les prédictions démontre la capacité du modèle à capturer les nuances qui déterminent si un vin est perçu comme léger ou puissant.

Tableau 1. Les métriques de performance utilisées sont le coefficient de détermination (R^2) et le coefficient de corrélation de Spearman (ρ) pour les prédictions sur l'ensemble de test, dérivées de modèles capables de prédire plusieurs caractéristiques (comme décrit dans la section Méthodes). Le tableau présente les valeurs de R^2 et ρ pour chaque modèle, appliquées à cinq caractéristiques distinctes du vin : Texture, Effervescence (pour les vins effervescents), Puissance (pour les vins rouges), Douceur et Structure (pour les vins rouges). Les meilleures performances pour chaque métrique sont indiquées en gras, soulignant les modèles les plus performants.

Modèle	Texture		Effervescence (vins effervescents)		Puissance (vins rouges)		Douceur		Structure (vins rouges)	
	R^2	ρ	R^2	ρ	R^2	ρ	R^2	ρ	R^2	ρ
Random Forest	0,125	0,338	0,444	0,633	0,765	0,683	0,411	0,634	0,712	0,743
Gradient Boosting	0,097	0,295	0,403	0,612	0,756	0,662	0,368	0,591	0,699	0,731
AdaBoost	0,013	0,186	0,462	0,635	0,684	0,640	0,171	0,503	0,620	0,713
Linear Regression	0,041	0,218	0,175	0,567	0,560	0,620	0,234	0,475	0,491	0,648
Lasso Regression	0,037	0,194	0,229	0,557	0,553	0,611	0,213	0,436	0,484	0,628
Support Vector Regression	0,010	0,158	0,199	0,390	0,306	0,557	0,179	0,408	0,386	0,576
Extra Trees	-0,832	0,110	-0,093	0,431	0,445	0,540	-0,263	0,383	0,327	0,566
XGBoost	0,017	0,259	0,342	0,585	0,743	0,662	0,357	0,589	0,685	0,720
CatBoost	0,110	0,325	0,453	0,636	0,785	0,686	0,416	0,634	0,737	0,744

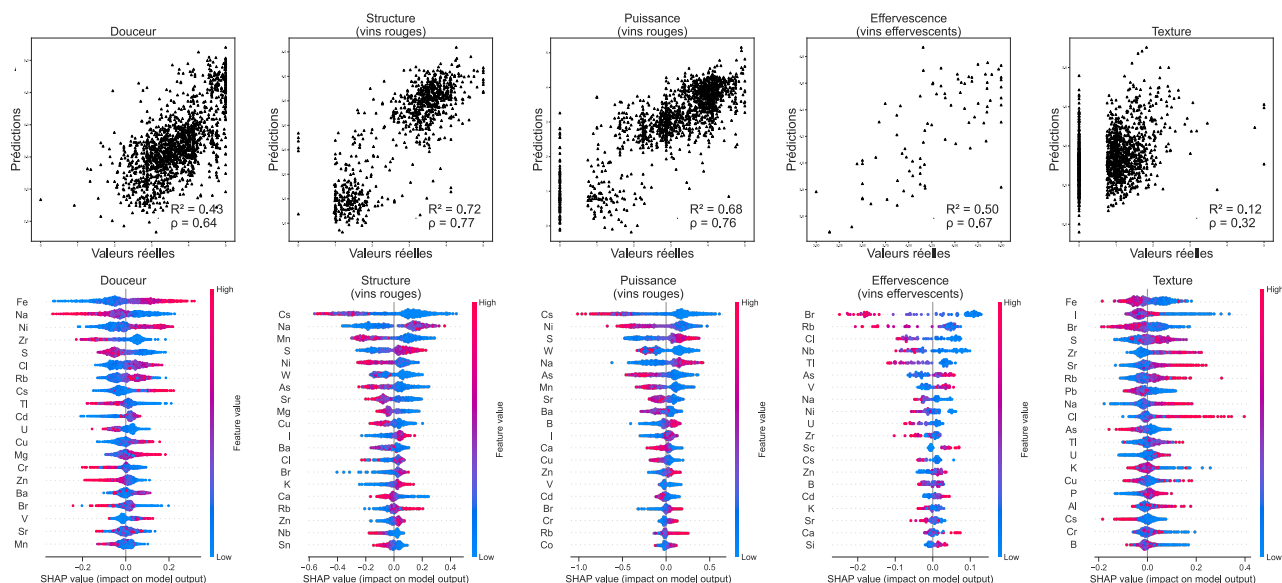


Figure 2. Évaluation des performances prédictives des modèles pour les caractéristiques structurales du vin (A) et analyse de l'importance des minéraux via les diagrammes SHAP (B). Le graphique présente les performances des modèles de prédiction pour cinq caractéristiques structurales du vin : Douceur, Structure (pour les vins rouges), Puissance (pour les vins rouges), Effervescence (pour les vins effervescents) et Texture. Pour chaque caractéristique, un graphique de dispersion montre la relation entre les valeurs réelles et les prédictions du modèle, accompagné des coefficients de détermination (R^2) et de corrélation de Spearman (ρ). En dessous de chaque graphique, un diagramme SHAP (SHapley Additive exPlanations) illustre l'importance relative des différents minéraux dans la prédiction de chaque caractéristique. Chaque point sur le diagramme SHAP représente un échantillon du jeu de données, avec la couleur indiquant la concentration du minéral (bleu pour les valeurs basses, rouge pour les valeurs élevées). Plus la valeur absolue sur l'axe horizontal est grande, plus l'impact du minéral sur la prédiction du modèle est significatif, indiquant les éléments les plus influents dans la modélisation des attributs organoleptiques du vin.

Structure (Souple à Tannique, vins rouges), $R^2 = 0,72$, $\rho = 0,77$: La prédiction de la structure, qui va d'une sensation souple à tannique, est particulièrement robuste. Le césium (Cs), le sodium (Na), et le manganèse (Mn) se distinguent comme des éléments clés influençant cette caractéristique. Ces minéraux sont associés à la formation et à la robustesse des tanins, contribuant ainsi à la sensation tannique dans les vins rouges. Le modèle capture efficacement l'effet de ces minéraux, ce qui souligne leur importance dans la perception de la structure du vin.

Effervescence (Tranquille à Mousseux, vins effervescents), $R^2 = 0,50$, $\rho = 0,67$: L'effervescence, une caractéristique allant de tranquille à mousseux, est bien prédite par le modèle. Des éléments comme le brome (Br), le rubidium (Rb), et le chlore (Cl) apparaissent comme particulièrement influents. Ces minéraux jouent un rôle direct dans la formation et la stabilité des bulles, expliquant ainsi leur contribution significative à la perception de l'effervescence dans les vins mousseux.

Douceur (Doux à Acide), $R^2 = 0,43$, $\rho = 0,64$: La douceur, qui est en partie modulée par l'acidité perçue,

montre des résultats de prédiction solides. Le fer (Fe), le sodium (Na), et le nickel (Ni) se révèlent exercer une influence notable sur cette caractéristique. Leur rôle est significatif dans la prédiction de la douceur, illustrant la complexité de cette perception sensorielle qui oscille entre doux et acide.

Texture (Sec à Moelleux), $R^2 = 0.12$, $\rho = 0.32$: La prédiction de la texture, qui va de sec à moelleux, est la moins précise, mais reste pertinente. Le fer (Fe), l'iode (I), et le brome (Br) ont un impact modéré sur la perception de la texture. Bien que ces minéraux influencent le ressenti en bouche, la texture est une caractéristique sensorielle complexe souvent dominée par des composés organiques et des interactions chimiques plus complexes, ce qui expliquerait pourquoi les résultats apparaissent moins robustes que pour d'autres attributs.

3.4. Analyse des caractéristiques aromatiques et saveurs

3.4.1. Sélection de l'algorithme de *Machine Learning*

Afin de garantir la stratégie de *machine learning* la plus adéquate, l'AUC de la classification binaire pour chaque descripteur organoleptique a été calculée. Cette opération a été répétée 5 fois pour chaque modèle, et les résultats moyens sont présentés dans le tableau 2. Sur l'ensemble des descripteurs organoleptiques étudiés, le modèle **CatBoost** présente une performance globale supérieure ou égale aux autres modèles testés. Avec des AUC élevées, en particulier pour les descripteurs essentiels comme **Puissant**, **Tannique**, et **Acide**. Le modèle **CatBoost** sera utilisé comme modèle principal.

Tableau 2. Performances des modèles de Machine Learning pour la classification des descripteurs organoleptiques du vin, évaluées par l'AUC moyenne. Ce tableau présente les AUC moyennes (Area Under the Curve) obtenues pour différents descripteurs organoleptiques du vin à l'aide de six modèles de *machine learning*.

Description organoleptique	Mean AUC						Number of sample	Number of target
	Random Forest	Logistic Regression	Gradient Boosting Machine	Conditional Random Forest	XGBoost	CatBoost		
Végétale	0,606	0,564	0,581	0,601	0,55	0,595	6275	974
Florale	0,633	0,631	0,636	0,632	0,633	0,65	6252	803
Fruits rouges ou noirs	0,825	0,8	0,816	0,827	0,815	0,824	7262	5117
Fruits blancs	0,716	0,707	0,717	0,722	0,69	0,721	7298	4558
Agrumes	0,857	0,847	0,851	0,861	0,853	0,858	7008	2384
Fruits exotiques	0,827	0,804	0,82	0,833	0,807	0,82	6657	895
Épices et aromates	0,672	0,659	0,667	0,67	0,645	0,672	6907	3463
Boisée	0,719	0,684	0,705	0,718	0,688	0,717	7401	3196
Secs	0,617	0,692	0,624	0,628	0,609	0,622	5984	581
Balsamique	0,681	0,687	0,692	0,669	0,727	0,708	7717	130
Empyreumatique	0,701	0,668	0,702	0,703	0,68	0,704	7125	2925
Animale	0,784	0,77	0,782	0,787	0,774	0,791	7509	1600
Minérale	0,725	0,714	0,725	0,729	0,704	0,729	7328	1929
Lactée	0,725	0,699	0,721	0,726	0,704	0,731	6125	2196
Pâtisserie	0,776	0,764	0,776	0,781	0,764	0,785	7476	1315
Amylique	0,725	0,716	0,743	0,744	0,732	0,776	7321	277
Défauts	0,683	0,656	0,681	0,69	0,656	0,686	7382	2119
Moelleux	0,696	0,641	0,684	0,7	0,672	0,707	4621	2310
Mousseux	0,938	0,934	0,925	0,941	0,93	0,941	282	141
Puissant	0,957	0,903	0,951	0,959	0,956	0,961	4903	2451
Acide	0,88	0,807	0,858	0,887	0,871	0,881	4903	2451
Tannique	0,954	0,922	0,95	0,951	0,955	0,956	2796	1398

Les différences entre les performances des modèles sont majoritairement dues à la nature des descripteurs organoleptiques. Certains, comme «Fruits rouges ou noirs» ou «Agrumes», sont souvent plus dominants et facilement identifiables, cela explique les scores AUC plus élevés, supérieur à 80. D'autres descripteurs, plus subtils moins dominants, comme «Végétale» ou «Boisée», présentent des AUC plus bas en raison de la subjectivité accrue de ces caractéristiques et de leur moindre présence dans les descriptions des vins analysés.

3.4.2. Analyse et explication des prédictions des données organoleptiques fruits rouges ou noirs.

Les graphiques (Fig. 3) ajoutent un niveau de précision et d'explicabilité aux résultats obtenus précédemment. L'objectif est de donner une vision plus approfondie et tangible des performances des modèles de *machine learning* et des facteurs influençant la prédiction des notes de fruits rouges ou noirs dans les vins, des analyses similaires ont été faites pour toutes les caractéristiques organoleptiques présente dans le tableau. Nous présentons les résultats pour les notes de fruits rouge ou noir. La courbe ROC (voir Fig. 3B) montre des résultats

relativement proches, au niveau des courbes et des scores AUC respectifs. Cette figure permet de vérifier visuellement la bonne répartition des données dans le jeu de test, avec des courbes lisses qui témoignent d'un grand nombre de points de données et d'une bonne répartition des classes (présence ou absence de fruits rouges ou noirs). La courbe ROC montre une performance bien équilibrée entre les différents modèles, ce qui renforce la robustesse des prévisions obtenues. L'axe "x" montre la proportion de négatifs incorrectement classés comme positifs (faux positifs), tandis que l'axe "y" montre la proportion de positifs correctement classés. Le principe de la courbe ROC est de mesurer la performance des modèles en fonction de différents seuils (thresholds). En changeant ce seuil, on peut ajuster le compromis entre la sensibilité et la spécificité du modèle. La sensibilité (True Positive Rate) est la capacité du modèle à détecter les vrais positifs, donc à bien prédire la présence d'une caractéristique. Plus elle est élevée, moins de vrais positifs sont oubliés dans les prédictions. La spécificité (1 - False Positive Rate) est la capacité du modèle à éviter les faux positifs, donc à ne pas attribuer une caractéristique à des échantillons qui ne la possèdent pas. En jouant sur le seuil, on peut ajuster la priorité du modèle selon l'objectif.

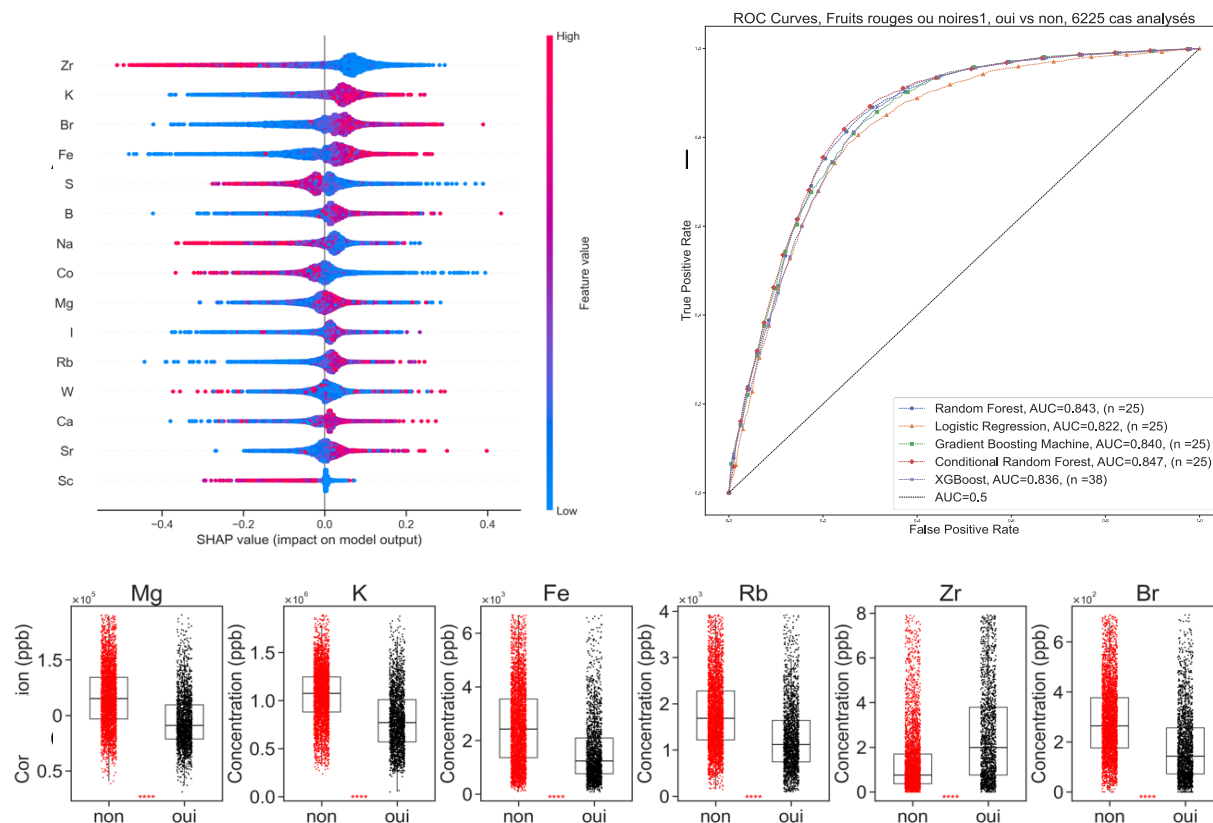


Figure 3. Évaluation de la pertinence des modèles de machine learning pour la prédiction de la présence de fruits rouges ou noirs dans les vins. Un diagramme SHAP montre l'impact relatif des différents éléments minéraux sur la prédiction de cette caractéristique (A). Les valeurs SHAP positives indiquent une plus grande probabilité de prédire la présence de fruits rouges ou noirs, tandis que les valeurs négatives suggèrent une absence. Les courbes ROC pour les modèles Random Forest, Régression Logistique, Gradient Boosting Machine, Forêt Aléatoire Conditionnelle, et XGBoost sont affichées (B). L'AUC (Area Under the Curve) pour chaque modèle est indiquée, montrant que la Forêt Aléatoire Conditionnelle (Conditional Random Forest) offre la meilleure performance pour cette classification avec une AUC de 0,847. Les boxplots illustrent les concentrations de certains minéraux clés (Mg, K, Fe, Rb, Zr, Br) comparées entre les vins présentant (oui) ou non (non) des notes de fruits rouges ou noirs (C). Ces graphiques montrent la distribution des concentrations en fonction de la présence ou l'absence de cette caractéristique sensorielle, soulignant l'importance de ces éléments dans la prédiction de cette note spécifique.

Dans cette étude, les diagrammes SHAP ont été utilisés pour approfondir la compréhension des paramètres qui influencent le plus la présence de fruits rouges ou noirs dans le vin (Fig. 3A). L'approche SHAP permet de décomposer les prédictions du modèle en attributs individuels et de quantifier la contribution de chaque paramètre à une prédiction donnée. L'analyse des contributions SHAP montre que certains éléments minéraux se distinguent particulièrement par leur influence sur la prédiction des fruits rouges ou noirs. Par exemple, le zirconium (Zr) apparaît comme un des minéraux les plus impactant dans la prédiction de cette caractéristique. Bien que le rôle du zirconium dans les profils organoleptiques du vin soit peu documenté, son importance pourrait être liée à des interactions complexes avec d'autres composés minéraux influençant la perception des arômes fruités et/ou une signature d'un procédé œnologique (utilisation de céramiques par exemple) qui aurait un impact indirect. De même, le potassium (K) et le fer (Fe) sont également des contributeurs clés, ces éléments étant historiquement associés à la maturation et à la structure des vins, influençant ainsi la perception de certains arômes.

Le diagramme SHAP révèle également que des éléments comme le sodium (Na) et le magnésium (Mg) ont un impact notable mais moindre, ce qui suggère qu'ils jouent un rôle secondaire dans les étapes de la structuration des arômes fruités. Ce type d'analyse permet de démystifier le fonctionnement du modèle et d'expliquer pourquoi certains vins sont prédits comme ayant des notes de fruits rouges ou noirs, en reliant directement ces prévisions à des propriétés chimiques spécifiques. Enfin, l'importance du rubidium (Rb) et du brome (Br), qui apparaissent dans le diagramme SHAP, soulève des questions intéressantes sur leur rôle potentiel dans les arômes des vins. Bien que ces éléments soient présents en traces, leur influence sur la structure aromatique pourrait provenir d'interactions spécifiques au niveau moléculaire, ou de leur corrélation et sensibilité plus grande avec d'autres composés plus dominants (respectivement potassium et chlore pour ces deux éléments). Ces résultats permettent de faire émerger des pistes d'investigation futures sur la relation entre les profils minéraux et les caractéristiques organoleptiques complexes du vin.

Les boxplots (Fig. 3C) illustrent la distribution des minéraux qui ressortent comme les plus différenciants pour la prédiction de la présence ou absence de fruits rouges ou noirs dans le vin. Ces éléments ont été identifiés comme significativement différents entre les deux groupes, avec des p-values inférieures à 0,0001, ce qui montre une distinction nette entre les vins analysés. Par exemple, on observe des écarts de concentration notables pour des éléments comme le magnésium (Mg) ou le fer (Fe), qui jouent probablement un rôle clé dans la formation des composés aromatiques du vin. Ces figures ajoutent

donc une couche supplémentaire d'analyse en montrant qu'il existe des différences minérales significatives entre les groupes, justifiant ainsi la performance des modèles

4. Conclusion

Cette étude montre que l'utilisation de l'IA associée aux profils minéraux des vins permet de prédire avec une bonne sensibilité et spécificité les qualités et caractéristiques organoleptiques principales des cuvées. L'une des perspectives majeures de notre approche réside dans la capacité de l'IA associée à ces données, à exploiter les jumeaux numériques des vins via les profils minéraux des différentes cuvées artificiellement générées pour tester une large gamme de combinaisons d'assemblages et trouver ainsi la meilleure pour répondre à une cible qualitative de production (Fig. 4).

Les profils minéraux reflètent la complexité des itinéraires vitivinicoles dans leur intégralité. Cette étude démontre que le profil minéral d'un vin est étroitement lié à ses caractéristiques organoleptiques et peut être utilisé, grâce à l'intelligence artificielle, pour anticiper sa qualité. Chaque cuvée possède un profil minéral unique, constitué de la concentration d'une quarantaine de minéraux exploitables. En utilisant ces informations, l'IA peut proposer des mélanges en ajustant les proportions de chaque cuvée, tout en respectant les limites des profils existants. Le processus fonctionne en ajustant les proportions de chaque vin choisi dans l'assemblage à des intervalles précis. Cela permet de simuler des centaines, voire des milliers de combinaisons, chacune ayant un profil minéral spécifique calculé en fonction des concentrations présentes dans chaque cuvée et des proportions dans lesquelles elles sont mélangées. Le producteur peut ainsi, non seulement identifier les assemblages qui se rapprochent le plus du profil minéral d'un vin souhaité. Que ce soit pour se rapprocher de vins de références, répondre aux attentes d'un panel spécifique de consommateurs ou simplement explorer des combinaisons pour influencer des caractéristiques organoleptiques spécifiques : des vins plus fruités, plus boisés, avec des tanins plus ou moins marqués, ou encore des niveaux d'acidité et de minéralité ajustés.

L'approche permet une maîtrise fine de l'assemblage des cuvées, tout en maintenant la cohérence du profil minéral dans les limites des cuvées disponibles. Cette technique est particulièrement utile pour les œnologues cherchant à reproduire des profils spécifiques d'année en année ou à innover en créant des cuvées aux caractéristiques sensorielles inédites. Au-delà de ces optimisations d'assemblages, l'IA combiné au MWP des vins deviendra également un outil au cœur des exploitations permettant d'identifier les modifications des processus viticoles et vinicoles pour ajuster les productions.

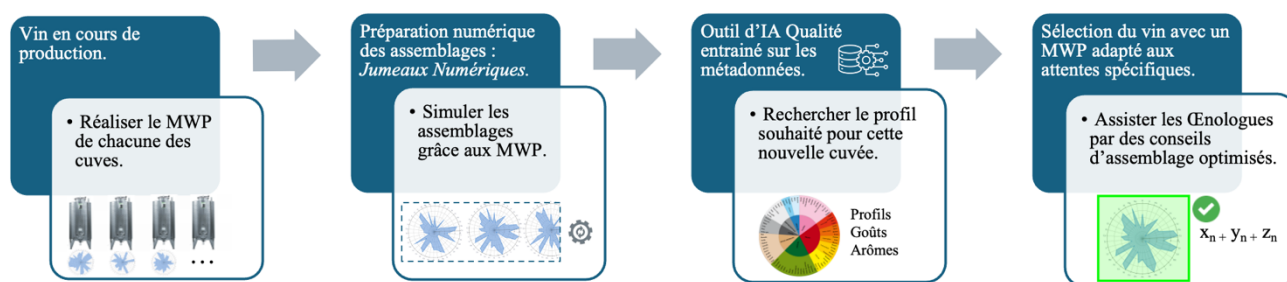


Figure 4. Objectif du développement d'un nouvel outil : IASSEMBLAGE, pour optimiser l'assemblage de cuvée en fonction d'un profil souhaité : Nouvelle cuvée ciblée / Expérience gustative maximale / Qualité globale optimale / Profil aromatique attendu / Satisfaction des consommateurs.

5. Références

1. G. Astray, C. Martinez-Castillo, J.-C. Mejuto, et J. Simal-Gandara, « Metal and metalloid profile as a fingerprint for traceability of wines under any Galician protected designation of origin », *J. Food Compos. Anal.*, vol. 102, p. 104043, sept. 2021
2. D. Tieman *et al.*, « A chemical genetic roadmap to improved tomato flavor », *Science*, vol. 355, n° 6323, p. 391-394, janv. 2017
3. M. Schartner, J. M. Beck, J. Laboyrie, L. Riquier, S. Marchand, et A. Pouget, « Predicting Bordeaux red wine origins and vintages from raw gas chromatograms », *Commun. Chem.*, vol. 6, n° 1, p. 1-10, déc. 2023
4. A. G. Potortí *et al.*, « Chemometric analysis of minerals and trace elements in Sicilian wines from two different grape cultivars », *Nat. Prod. Res.*, vol. 31, n° 9, p. 1000-1005, mai 2017
5. A. Loutfi, S. Coradeschi, G. K. Mani, P. Shankar, et J. B. B. Rayappan, « Electronic noses for food quality: A review », *J. Food Eng.*, vol. 144, p. 103-111, janv. 2015,
6. C. Villano *et al.*, « Wine varietal authentication based on phenolics, volatiles and DNA markers: State of the art, perspectives and drawbacks », *Food Control*, vol. 80, p. 1-10, oct. 2017
7. P. Pfister *et al.*, « Odorant Receptor Inhibition Is Fundamental to Odor Encoding », *Curr. Biol. CB*, vol. 30, n° 13, p. 2574-2587.e6, juill. 2020
8. Yohann Clément, Pierre Lantéri, Bertrand Chatelet, Théodore Tillement³, Coraline Duroux, Leticia Gomes, Olivier Tillement, « Identification spécifique des dix crus du beaujolais »
9. L. Sarlo, C. Duroux, Y. Clément, P. Lanteri, F. Rosseti, O. David, A. Tillement, P. Gillet, A. Hagège, L. David, M. Dumoulin, R. Marchal, T. Tillement, F. Lux, and O. Tillement, OENO One., « Enhancing Wine Authentication: Leveraging 12,000+ International Mineral Wine Profiles and Artificial Intelligence for Accurate Origin and Variety Prediction ». Oeno one, (en préparation).
10. P. A. Gonzalez *et al.*, « Elemental composition, rare earths and minority elements in organic and conventional wines from volcanic areas: The Canary Islands (Spain) », *PLOS ONE*, vol. 16, n° 11, p. e0258739, nov. 2021
11. M. Perini et L. Bontempo, « Liquid Chromatography coupled to Isotope Ratio Mass Spectrometry (LC-IRMS): A review », *TrAC Trends Anal. Chem.*, vol. 147, p. 116515, févr. 2022
12. D. Bertoldi, R. Larcher, M. Bertamini, S. Otto, G. Concheri, et G. Nicolini, « Accumulation and Distribution Pattern of Macro- and Microelements and Trace Elements in *Vitis vinifera* L. cv. Chardonnay Berries », *J. Agric. Food Chem.*, vol. 59, p. 7224-36, juin 2011
13. Coraline Duroux, Théodore Tillement, Pierre Der Nigohossian, Agnès Hagège, Laurent David, Michel Dumoulin, Olivier Tillement, « La Revue des Œnologues | Analyse multiélémentaire des vins : « Metal Wine Profile » »
14. P. Yu, M. Y. Low, et W. Zhou, « Design of experiments and regression modelling in food flavour and sensory analysis: A review », *Trends Food Sci. Technol.*, vol. 71, p. 202-215, janv. 2018
15. M. Malavolta *et al.*, « A survey on computational taste predictors », *Eur. Food Res. Technol.*, vol. 248, n° 9, p. 2215-2235, sept. 2022
16. H. W. Moskowitz, V. Kumaraiah, K. N. Sharma, H. L. Jacobs, et S. D. Sharma, « Cross-cultural differences in simple taste preferences », *Science*, vol. 190, n° 4220, p. 1217-1218, déc. 1975
17. G. M. Shepherd, « Smell images and the flavour system in the human brain », *Nature*, vol. 444, n° 7117, p. 316-321, nov. 2006
18. F. Fritz, R. Preissner, et P. Banerjee, « VirtualTaste: a web server for the prediction of organoleptic properties of chemical compounds », *Nucleic Acids Res.*, vol. 49, n° W1, p. W679-W684, juill. 2021
19. O. Oladokun, A. Tarrega, S. James, K. Smart, J. Hort, et D. Cook, « The impact of hop bitter acid and polyphenol profiles on the perceived bitterness of beer », *Food Chem.*, vol. 205, p. 212-220, août 2016

20. B. K. Lee *et al.*, « A principal odor map unifies diverse tasks in olfactory perception », *Science*, vol. 381, n° 6661, p. 999-1006, sept. 2023
21. L. Zhang, Q. Liu, Y. Li, S. Liu, Q. Tu, et C. Yuan, « Characterization of wine volatile compounds from different regions and varieties by HS-SPME/GC-MS coupled with chemometrics », *Curr. Res. Food Sci.*, vol. 6, p. 100418, déc. 2022
22. E. Chambers et K. Koppel, « Associations of volatile compounds with sensory aroma and flavor: the complex nature of flavor », *Mol. Basel Switz.*, vol. 18, n° 5, p. 4887-4905, avr. 2013
23. N. Eriksson *et al.*, « A genetic variant near olfactory receptor genes influences cilantro preference », *Flavour*, vol. 1, n° 1, p. 22, nov. 2012
24. [S. Chinchanchokchai, P. Thontirawong, et P. Chinchanchokchai, « A tale of two recommender systems: The moderating role of consumer expertise on artificial intelligence based product recommendations », *J. Retail. Consum. Serv.*, vol. 61, p. 102528, juill. 2021
25. D. J. Cook, T. A. Hollowood, R. S. T. Linforth, et A. J. Taylor, « Correlating instrumental measurements of texture and flavour release with human perception », *Int. J. Food Sci. Technol.*, vol. 40, n° 6, p. 631-641, 2005
26. G. P. Danezis, A. S. Tsagkaris, V. Brusic, et C. A. Georgiou, « Food authentication: state of the art and prospects », *Curr. Opin. Food Sci.*, vol. 10, p. 22-31, août 2016
27. A. Legin, A. Rudnitskaya, B. Seleznev, et Yu. Vlasov, « Electronic tongue for quality assessment of ethanol, vodka and eau-de-vie », *Anal. Chim. Acta*, vol. 534, n° 1, p. 129-135, avr. 2005
28. M. Kermit et V. Lengard, « Assessing the performance of a sensory panel–panellist monitoring and tracking », *J. Chemom.*, vol. 19, n° 3, p. 154-161, mars 2005
29. B. Plutowska et W. Wardencki, « Application of gas chromatography–olfactometry (GC–O) in analysis and quality assessment of alcoholic beverages – A review », *Food Chem.*, vol. 107, n° 1, p. 449-463, mars 2008
30. S. Catarino, A. S. Curvelo-Garcia, et R. B. de Sousa, « Measurements of contaminant elements of wines by inductively coupled plasma-mass spectrometry: A comparison of two calibration approaches », *Talanta*, vol. 70, n° 5, p. 1073-1080, déc. 2006
31. M. Schreurs *et al.*, « Predicting and improving complex beer flavor through machine learning », *Nat. Commun.*, vol. 15, n° 1, p. 2368, mars 2024
32. N. Costa, J. P. Ximenez, J. Rodrigues, F. Barbosa, et R. Barbosa, « Characterization of Cabernet Sauvignon wines from California: determination of origin based on ICP-MS analysis and machine learning techniques », *Eur. Food Res. Technol.*, vol. 246, juin 2020
33. T. Fawcett, « An introduction to ROC analysis », *Pattern Recognit. Lett.*, vol. 27, n° 8, p. 861-874, juin 2006
34. P. Kment, M. Mihaljevič, V. Ettler, O. Šebek, L. Strnad, et L. Rohlová, « Differentiation of Czech wines using multielement composition – A comparison with vineyard soil », *Food Chem.*, vol. 91, p. 157-165, juin 2005