



**HAL**  
open science

# Typologie de l'arrachage viticole héraultais en 2007 par Classification Hiérarchique Ascendante

Samson Zadmeهران, Etienne Montaigne

## ► To cite this version:

Samson Zadmeهران, Etienne Montaigne. Typologie de l'arrachage viticole héraultais en 2007 par Classification Hiérarchique Ascendante. 25. Oenométrie, Vineyard Data Quantification Society (VDQS). INT., May 2018, Dijon, France. hal-02736548

**HAL Id: hal-02736548**

**<https://hal.inrae.fr/hal-02736548>**

Submitted on 2 Jun 2020

**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.



***EUAWE DIJON, 23-25 mai 2018***

## **Typologie de l'arrachage viticole héraultais en 2007 par Classification Hiérarchique Ascendante.**

**Samson Zadmehrán**, Doctorant, Montpellier Université

**Etienne Montaigne**, Professeur Émérite, Montpellier SupAgro

### **Résumé :**

Dans cet article nous prolongeons notre précédente étude : « La politique d'arrachage a-t-elle répondu à ses objectifs en 2007 : Une étude de cas sur un échantillon départemental » (Montaigne et al., 2017). Nous repartons des mêmes données issues de 341 dossiers Viniflor d'arrachage et de 50 entretiens qualitatifs. Nous avons recours cette fois-ci aux méthodes plus puissantes de l'analyse des données : l'analyse en composante principale associée à la classification hiérarchique ascendante. Ces méthodes nous apportent un complément d'analyse statistique en vue de valider les résultats précédents. Dans un premier temps, nous utilisons des méthodes classiques d'analyses de données, mais celles-ci manquent d'efficacité à cause des deux niveaux des données — entreprise et parcelle — et de la disproportion entre les nombres de certaines variables par rapport à d'autres. Pour résoudre ce problème, nous avons eu recours à des jointures, c'est-à-dire la création de plusieurs tables de données combinant des données initiales venant des deux niveaux. Nous appliquons grâce à cela successivement les deux méthodes, elles-mêmes complétées par un test de Chi<sup>2</sup>. Ceci nous permet de créer deux séries de clusters, l'une au niveau de l'entreprise et l'autre niveau des parcelles. Ces traitements nous permettent de conclure que nos nouveaux résultats n'infirment pas qualitativement les anciens, mais modifient significativement les effectifs.

## 1 Introduction

Lors de la précédente conférence — Bologne EUAWE Oenométrie 24 — nous avons présenté un article intitulé « La politique d'arrachage a-t-elle répondu à ses objectifs en 2007 : Une étude de cas sur un échantillon départemental » (Montaigne et al., 2017). Cet article visait à évaluer l'impact de la politique économique des arrachages menée par la Commission européenne dans le cadre de l'OCM vin 2008 anticipée. Dans cet article, notre étude se limite à un département : l'Hérault. Nos résultats ne visent pas à donner une interprétation globale de cette politique d'arrachage à l'échelle de l'ensemble des pays de l'Union européenne. Elle vise modestement à comparer dans l'Hérault les résultats ciblés et énoncés par la Commission européenne aux résultats obtenus dans le département.

L'un des points majeurs de l'OCM vin 2008 était la programmation de l'élimination entre 2015 et 2018 des droits de plantations qui régulent le potentiel de production du vignoble européen (Delord et al., 2015). Cette disparition aurait permis l'avènement de la libéralisation du secteur viticole européen qui, en théorie et selon les arguments du moment, aurait favorisé la compétitivité des vignerons européens à l'international (European Commission, 2006). Avant cette libéralisation, une autre politique visant à améliorer la compétitivité de l'ensemble des vignerons devait être menée : les campagnes d'arrachages primées (Montaigne et al., 2017). Le but était d'éliminer les vignerons les moins productifs grâce à trois séries de campagnes d'arrachage, 2008-2009, 2009-2010 et 2010-2011. La Commission européenne avait prévu un budget permettant l'arrachage de 175 000 hectares de vignes (Batzeli, 2007). Au final, cette campagne d'arrachage a été un franc succès, car l'ensemble des 175 000 hectares primables a été arraché.

Les arrachages ne sont pas obligatoires, se font sur la base du volontariat et sont primés. La Commission européenne donne la priorité aux dossiers de ceux qui arrachent l'intégralité de leurs vignes et qui sont âgés de plus de 55 ans. Les primes augmentent avec le rendement des vignes arrachées. Les viticulteurs souhaitant arracher leurs vignes ont dû compléter un dossier qui, en France, était soumis à Viniflor. L'accès à ces dossiers est normalement restreint. Cependant, à la demande des professionnels, chaque vigneron concerné pouvait autoriser le partage de son dossier avec les instituts de recherche et les organismes de développement. Ce fut le cas dans le département de l'Hérault, où la Chambre d'Agriculture a mis à notre disposition 341 dossiers d'arrachage. Ceux-ci représentaient près de 20 % du total des dossiers dans l'Hérault, soit 484 hectares de vignes arrachées.

Dans notre communication présentée à Bologne, les résultats abondaient principalement de l'étude des dossiers à l'aide de statistiques descriptives. Nous avons d'abord mis en relation, par dossier d'exploitation viticole, les pourcentages de vignes arrachées avec la superficie totale en vigne. Ceci nous a permis d'isoler 4 groupes. Le groupe 1 est constitué de petites exploitations viticoles (<10 ha) arrachant près de 100 % de leurs vignes, quittant ainsi définitivement la viticulture. Le groupe 2 inclut des entreprises viticoles de taille similaire, mais qui arrachent à un taux compris entre 80 % et 50 %. Les exploitations du groupe 2 gardent donc quelques parcelles de vigne. Notre interprétation était que ces deux groupes correspondaient aux cibles désignées par la Commission européenne.

Nous avons pu déterminer la présence de deux autres groupes. Le groupe 3 est constitué d'entreprises arrachant moins de 50 % de leur superficie plantée en vigne. Leur taille peut

atteindre jusqu'à 25 ha. Finalement, le groupe 4 est constitué d'exploitations viticoles dont la taille varie entre 25 ha et 80 ha. Leur taux d'arrachage est faible, mais la valeur des surfaces arrachées est forte.

Afin de cerner les motivations de ces arrachages, nous avons effectué une enquête qualitative supplémentaire sur une cinquantaine de vigneron choisis parmi les 341 dossiers. Les résultats de cette enquête montrent que les motivations principales oscillent entre des raisons de cycles de vie du viticulteur (retraite) et des problèmes de trésorerie liés aux crises successives sur le marché du vin (2004 : crise de surproduction mondiale, 2007 : crise des *subprimes* – crise de la demande) (Centre d'Économie Rurale (CER) France, 2009 ; Centre d'Économie Rurale (CER) Languedoc-Roussillon, 2006). Mais d'autres raisons telles que la possibilité de vendre des terres (libérées après les arrachages) en temps de crise financière se manifestent. Pour le groupe 4, arracher des vignes pouvait constituer aussi une valeur d'option, c'est-à-dire disposer d'une trésorerie permettant de reporter la décision de continuer ou non l'activité viticole.

Les dossiers contenaient principalement des données sur les cépages et les âges des vignes arrachées. Notre analyse faisait apparaître que 1/6 des vignes arrachées ne correspondaient pas, a priori, aux vignes ciblées par la Commission européenne. En effet elles étaient soit trop jeunes (< 25 ans), soit complantées en « cépages améliorateurs » (i.e. Merlot, Cabernet-Sauvignon...) (Montaigne et al., 2017).

En simplifiant nos résultats, nous pouvions dire que, pour moitié, la politique d'arrachage a atteint ses objectifs de « rajeunissement » du vignoble et des viticulteurs et pour moitié a permis à de nombreuses exploitations de « survivre » à la crise dans l'attente de jours meilleurs ou d'arrêt définitif de l'activité.

Lors de l'écriture de l'article précédent, nous avons en perspective d'améliorer les traitements statistiques de nos données en leur appliquant une classification hiérarchique ascendante (CAH) selon la méthode de Ward. Ce type de méthode pouvait nous apporter une typologie plus fine des viticulteurs en mobilisant dans l'analyse un nombre de variables plus important. Cependant sous le logiciel SPSS, nous avons rencontré un problème avec ce traitement situé au niveau des gains d'inertie. En effet, pour qu'une CAH soit représentative, ce sont les premières classes du dendrogramme qui doivent contenir le plus d'inertie (Montaigne et al., 2017). Sur cette première tentative, nous obtenions des résultats inverses (inertie maximale dans les dernières classes). Il était donc impossible d'étendre notre traitement statistique afin de contenir dans une analyse quantitative unique les informations issues des 4 groupes arrachant des vignes, leurs motivations, les cépages et les âges des vignes arrachées.

Grâce aux outils développés sous le logiciel R nous sommes aujourd'hui en mesure de franchir cet obstacle. Les bibliothèques programmées dans R, « FactoMineR » (Lê et al., 2008) et « Facto extra » (Kassambara, 2017), nous mettent à disposition tous les outils nécessaires pour réaliser cette CAH (Husson et al., 2016, p. 214). La méthode invite à concaténer les données dans une base de données globale (Cornillon et al., 2012, p. 43). Cette base fait l'objet d'un premier traitement selon la méthode de l'analyse des composantes principales (ACP) et qui peut être suivi par une CAH. Il est alors possible de présenter le dendrogramme

de la CAH en trois dimensions sur le premier plan factoriel de l'ACP (Husson et al., 2016, p. 194). Les données qualitatives issues du questionnaire sur les motivations des arrachages nous guideront dans l'interprétation (Husson et al., 2016, p. 21). Nos attentes sont ainsi de retrouver dans le dendrogramme de la CAH des groupes mieux caractérisés que par les statistiques descriptives à deux variables. L'association du premier plan factoriel de l'ACP et des variables qualitatives nous aidera dans l'analyse des clusters issus du dendrogramme. Les interprétations résultant de ce nouveau traitement statistique viendront compléter l'évaluation que nous avons émise sur la politique économique des arrachages de la Commission européenne dans l'Hérault.

## 2 Méthodologie

### 2.1 Traitement statistique ciblé

« L'objectif d'une Analyse en Composante Principale est de résumer un tableau de données individus x variables. L'ACP permet d'étudier les ressemblances entre individus du point de vue de l'ensemble des variables et dégage des profils d'individus. Elle permet également de réaliser un bilan des liaisons linéaires entre variables cadres, elles peuvent être reliées, ce qui permet de caractériser les individus ou groupes d'individus par les variables et d'illustrer les liaisons entre variables à partir d'individus caractéristiques. » (Cornillon et al., 2012, p. 201)

« La Classification Ascendante Hiérarchique a pour objectif de construire une hiérarchie sur les individus et se présente sous la forme d'un dendrogramme. Cette classification permet de regrouper des individus dans des classes les plus homogènes possibles à partir d'un jeu de données individus x variables. Cette méthode permet de choisir une distance entre les individus ou, plus généralement, une dissimilarité et un indice d'agrégation. Il est souvent recommandé d'utiliser l'indice d'agrégation moyen ou l'indice d'agrégation Ward. » (Cornillon et al., 2012, p. 230)

### 2.2 La CAH et Ward.

Selon Husson, l'« enchaînement des deux démarches fournit une possibilité méthodologique nouvelle : ne conserver pour la CAH qu'une partie des facteurs d'inertie non nulle. Pour cela, les deux raisonnements suivants peuvent être tenus.

– Eliminer de la CAH les seules dimensions dont on est (pratiquement) sûr qu'elles ne représentent que du "bruit" c'est-à-dire les toutes dernières ; ainsi on conservera les facteurs totalisant un pourcentage très élevé de l'inertie, disons 80 % ou 90 % pour fixer les idées ; la hiérarchie ainsi obtenue est censée être plus stable et plus claire.

– Ne conserver pour la CAH que les axes que l'on a su interpréter soit, en pratique, un tout petit nombre (entre 2 et 5 pour fixer les idées) ; la hiérarchie ainsi obtenue joue essentiellement le rôle d'aide à l'interprétation de l'analyse factorielle » (Husson et al., 2016, p. 188).

### **2.3 Les données : leur particularité.**

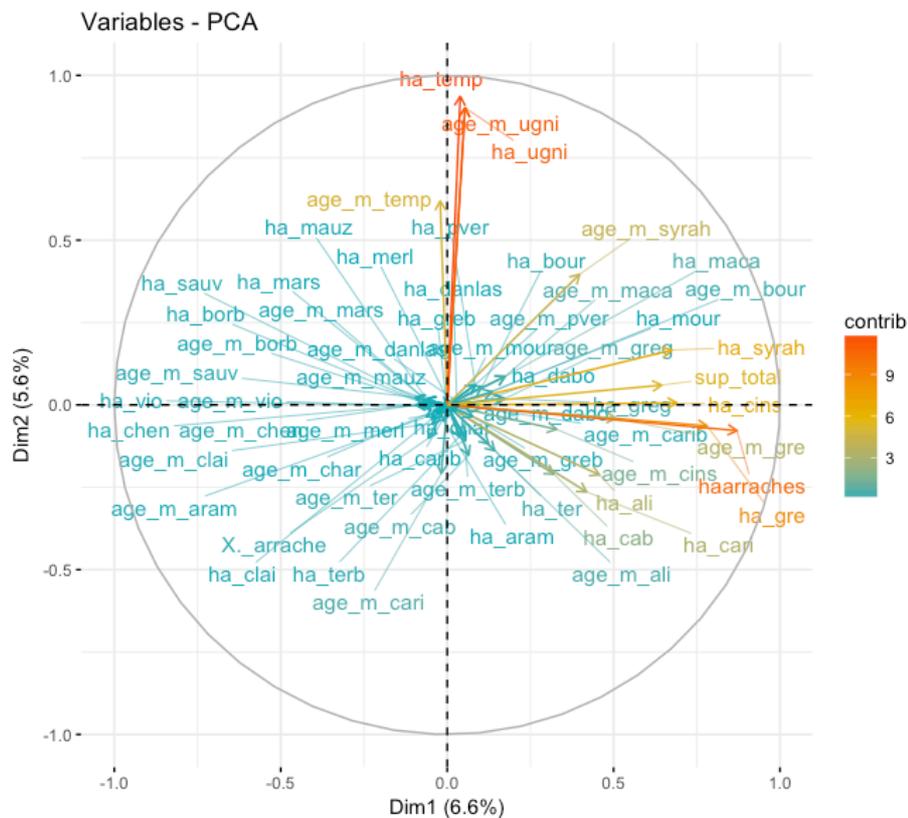
Les données que nous avons recueillies sont issues de deux sources. Les dossiers Viniflohr nous donnent les informations relatives à l'identité de l'entreprise, sa localisation géographique, l'identité cadastrale des parcelles détenant les vignes arrachées, la surface totale de l'entreprise viticole et, par parcelle, les cépages, les surfaces et l'âge des vignes arrachées. Parmi ces 341 dossiers, nous avons sélectionné 50 entreprises afin d'effectuer avec elles des entretiens qualitatifs et ainsi mieux comprendre les motivations de leurs arrachages.

La particularité de nos données est qu'elles caractérisent l'entreprise viticole à deux niveaux différents : certaines variables telles que la taille de l'entreprise viticole, décrivent l'entreprise à son niveau global alors que d'autres telles que l'âge et la nature des cépages arrachés l'étudient au niveau parcellaire. Les variables sont aussi de natures différentes, qualitatives pour les types de cépages, les raisons de l'arrachage et quantitatives pour les surfaces des parcelles et des entreprises, les taux d'arrachages et les âges des parcelles.

### **2.4 Structuration des données et premier échec.**

Notre première tentative a été de créer un unique fichier Excel regroupant par entreprise viticole l'intégralité des données pour les analyser dans leur ensemble. Ainsi nous avons défini une première série de variables au niveau de l'entreprise viticole résumant la surface totale de l'entreprise viticole, la surface totale arrachée, le pourcentage de surface arrachée. Nous avons introduit une seconde série de variables au niveau de la parcelle décrivant pour chacune d'elles la nature et l'âge du cépage arraché et la surface de la parcelle. Pour cela nous avons calculé pour chaque cépage : (1) la moyenne pondérée de l'âge des cépages arrachés en fonction de la taille de parcelle et (2) la somme des parcelles par entreprise viticole ayant un cépage identique.

Figure 1 : Cercle de corrélation des variables : les surfaces de chaque cépage arraché par exploitation et les moyennes d'âge de ces parcelles pour chaque cépage dans ces exploitations.



La conséquence de cette structure de tableau a été une présence disproportionnée des variables décrivant les cépages et les âges des vignes arrachées par rapport aux variables situées au niveau de l'entreprise viticole. En effet l'ACP cherche les axes factoriels décrivant la dispersion maximale des points représentant les entreprises selon les variables qui les décrivent. La disproportion entre d'une part les 3 variables traitant de la taille de l'entreprise viticole (la surface de l'entreprise, la surface arrachée et le pourcentage de surface arrachée) et d'autre part, les 60 variables — 30 décrivant l'âge pondéré de chaque cépage arraché et 30 donnant la superficie des parcelles ayant de vignes arrachées — rendent l'ACP inefficace (Figure 1). Dans la description des axes factoriels, les 3 variables situées au niveau des entreprises viticoles sont noyées parmi les 60 variables du niveau parcellaire traitant des cépages. En outre, utiliser la moyenne pondérée de l'âge des cépages arrachés en fonction de la taille de la parcelle, même si elle permet de réaliser l'analyse à partir d'un seul fichier, fait disparaître une dimension essentielle de l'évaluation de politique économique, à savoir l'âge de chaque parcelle arrachée. De façon assez logique, nos premières tentatives avec le package « FactoMineR » ont été insatisfaisantes, car la structure des données n'était pas en adéquation avec la méthode statistique employée.

En ce qui concerne la CAH, d'autres difficultés sont apparues. Une CAH cherche à rapprocher les membres dont les données sont semblables et à éloigner les membres avec des données différentes. Il est donc contre-productif d'introduire des moyennes pondérées d'âge, car cet usage réduit les différences et va l'encontre de l'utilisation d'une CAH. N'étant pas satisfaits de ces premiers résultats, nous avons modifié notre méthodologie.

## 2.5 Choix final de la structuration des données et de la séquence de traitement.

Face à ces problèmes, nous avons fait évoluer la méthode et structuré différemment les données. Nous avons gardé les 2 tableaux de données séparés : un premier tableau réunit les données d'entreprise et un second tableau les données parcellaires. Ces deux tableaux sont regroupés dans une base de données de type MySQL. Nous y ajoutons un troisième tableau regroupant les 50 entretiens qualitatifs.

La « clé » pour communiquer entre les trois tableaux est le numéro de dossier Viniflor inséré dans chaque tableau. Cette clé nous permet d'anonymiser les données et surtout de créer des sous-tableaux par les fonctions de jointure de R, tableaux que nous analyserons ensuite grâce aux traitements successifs de l'ACP et de l'CAH.

Nous donc deux sous-tableaux : l'un au niveau des entreprises, l'autre au niveau des parcelles. Le premier sous-tableau au niveau de l'entreprise permet d'étudier la relation entre deux groupes de variables : (1) les variables quantitatives de l'entreprise viticole : sa surface totale, sa surface arrachée et son pourcentage de surface arrachée et (2) les variables quantitatives des parcelles arrachées ventilées par classe d'âge et surfaces arrachées. Pour ce faire, à partir de la base de données des parcelles, pour chaque entreprise nous créons quatre classes de surface arrachée selon les intervalles d'âge de la vigne : 0-15, 16-30, 31-45, et 46+ ans. Nous éliminons les entreprises qui ont des données manquantes. Puis nous lançons le package de R « FactoMineR » pour effectuer le traitement statistique, suivi de « Facto extra » pour améliorer les résultats graphiques (Kassambara, 2017).

Le second sous-tableau est construit à partir d'une entrée « parcelles ». Les rangées sont les identifiants des parcelles. Les variables sont 1) l'identifiant Viniflor de l'entreprise viticole à laquelle elle appartient, (2) le cépage arraché sur la parcelle, (3) l'âge du cépage arraché, (4) la surface de la parcelle, (5) la surface totale en vigne de l'entreprise agricole et (6) la surface totale arrachée par cette entreprise viticole. À nouveau, nous utilisons les packages de R « FactoMineR » et « Facto extra » pour analyser et illustrer les données, les plans factoriels et les clusters.

À l'issue de ces deux traitements ACP et CAH, nous disposons de deux séries de clusters, le premier nous offre une typologie des entreprises et le second une typologie des parcelles. Pour la typologie des entreprises, nous utiliserons les données qualitatives issues des 50 entretiens qualitatifs pour essayer d'affiner notre interprétation au niveau des motivations d'arrachage. L'étude des clusters « parcelles » sera enrichie par un test de chi 2 sur la nature des cépages arrachés.

### 3 Interprétation du premier tableau issu d'une jointure par entreprise viticole

#### 3.1 ACP : les variables et le premier plan factoriel.

Le traitement « FactoMineR » nous donne deux résultats graphiques : un cercle de corrélations des variables de rayon 1 (Figure 2) (Bialès and Fenneteau, 1998) et un premier plan factoriel avec l'ensemble des individus (

À l'aide de ces premières informations, nous pouvons déjà donner une interprétation au placement des entreprises viticoles dans le premier plan factoriel (Figure 3). La valeur absolue des hectares de vignes augmente de gauche à droite. Le pourcentage des vignes arraché par rapport à la superficie moyenne augmente de bas en haut. La valeur des superficies des entreprises viticole augmente selon la diagonale allant d'en haut à gauche à en bas à droite. Nous étudierons la contribution des âges des vignes arrachées avec la CAH.

Figure 3), dans notre cas les entreprises viticoles. Nous nous intéressons au cercle qui nous permet de « visualiser les coefficients de corrélation entre les variables et les composantes principales » dans notre cas nous nous concentrons sur la première et la seconde dimension (Husson et al., 2016, p 12). La première dimension explique 39,5 % de l'inertie totale l'ACP et la seconde dimension explique 19,7 %. Cette information est visible dans la Figure 2 et le Tableau 1.

En regardant la Figure 2, on observe que trois variables sont proches du cercle de rayon 1. Ces trois variables sont très bien représentées dans le premier plan factoriel qui décrit les deux premières dimensions de l'ACP (Bialès & Fenneteau, 1998). La variable « haarraches » qui exprime la valeur totale en hectare de vignes arrachées est celle qui est la mieux représentée et elle est corrélée à 95 % à la dimension 1 selon le Tableau 1. Elle exprime environ 1/3 du premier plan factoriel. La variable « sup\_total » qui représente la superficie totale des entreprises viticoles contribue à 16 % de la construction de l'axe et est corrélée à 66 % avec l'axe 1. Ces deux variables contribuent à presque 50 % de la dimension 1.

La dimension 2 est, quant à elle, surtout marquée par deux variables : « sup\_total », la superficie totale de l'entreprise viticole, et « pour\_arr », le pourcentage d'arrachage de l'entreprise viticole. Elles contribuent à elles deux à 76 % de la construction de la seconde dimension. Le pourcentage d'arrachage est très fortement positivement corrélé à la seconde dimension (0,86) et la superficie totale est négativement corrélée avec la deuxième dimension (-0,55).

Figure 2 : cercle des corrélations de variables

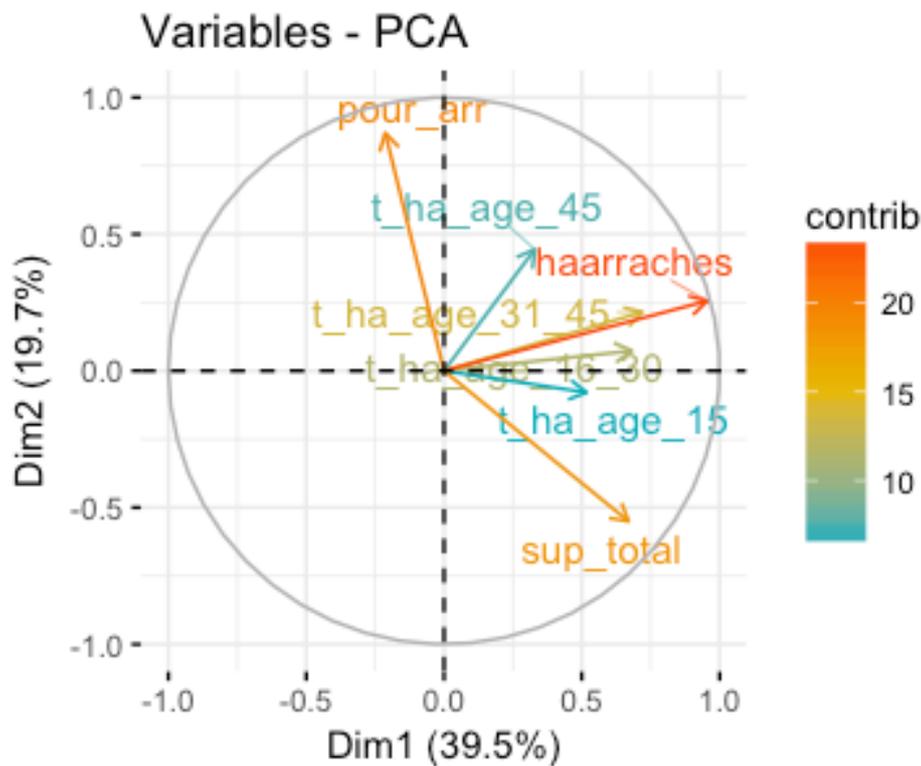
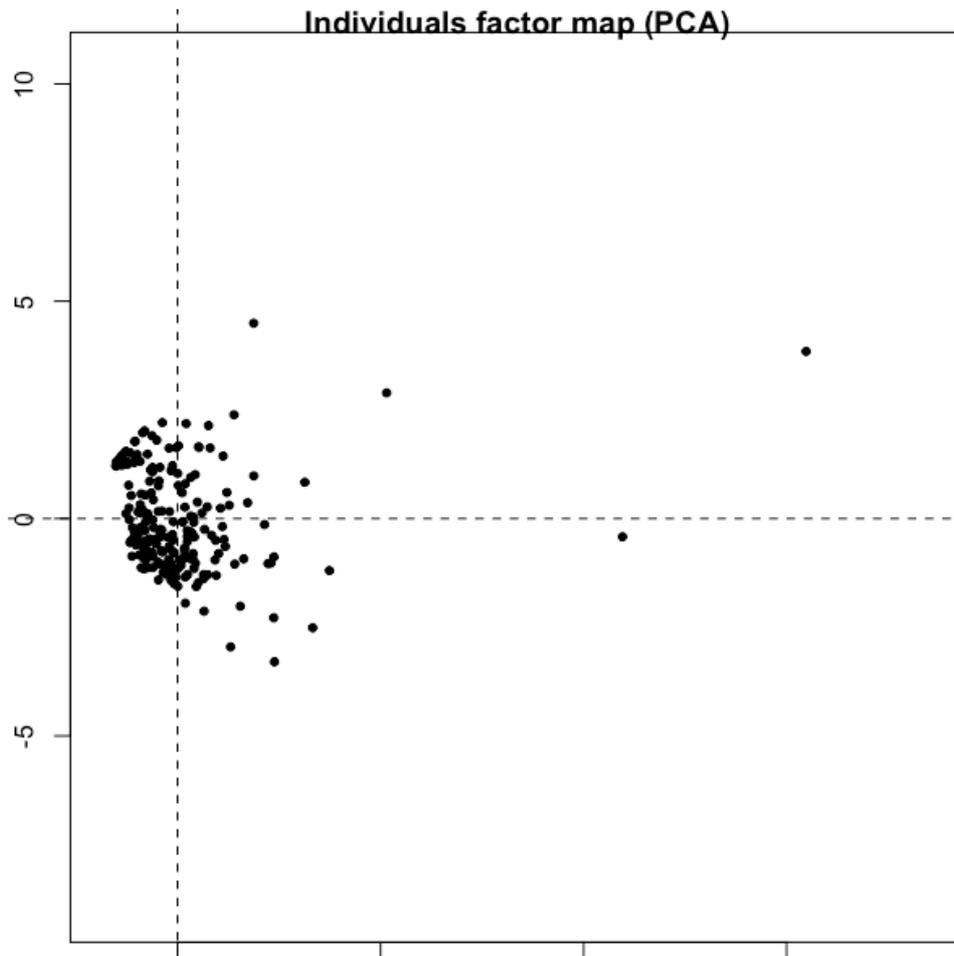


Tableau 1 : corrélation et contribution des variables aux 2 premières dimensions du plan factoriel

Variables	Dim.1 corrélacion	Dim.2 corrélacion	Dim.1 contribution	Dim.2 contribution
t_ha_age_15	0.517040436	-0.079112277	9.671701481	0.453636493
t_ha_age_16_30	0.681818344	0.073056729	16.81865366	0.386848279
t_ha_age_31_45	0.719975435	0.214845594	18.75379751	3.345593141
t_ha_age_45	0.329108616	0.444067086	3.918612947	14.29280243
haarraches	0.954689562	0.255750679	32.97449959	4.740823762
sup_total	0.669479082	-0.550094656	16.21540832	21.9328502
pour_arr	-0.21338452	0.869897521	1.647326496	54.8474457

À l'aide de ces premières informations, nous pouvons déjà donner une interprétation au placement des entreprises viticoles dans le premier plan factoriel (Figure 3). La valeur absolue des hectares de vignes augmente de gauche à droite. Le pourcentage des vignes arraché par rapport à la superficie moyenne augmente de bas en haut. La valeur des superficies des entreprises viticole augmente selon la diagonale allant d'en haut à gauche à en bas à droite. Nous étudierons la contribution des âges des vignes arrachées avec la CAH.

Figure 3 : Entreprise viticole dans le premier plan factoriel de l'ACP.



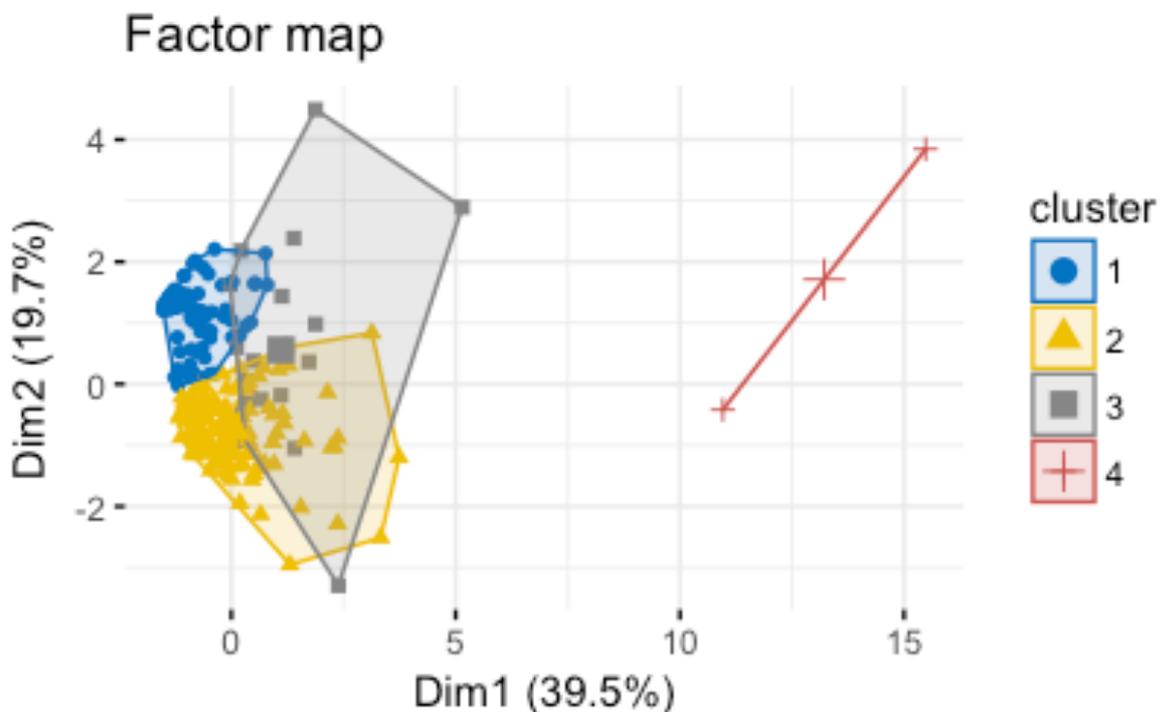
### 3.2 Classification hiérarchique ascendante.

À nouveau nous utilisons les packages « FactoMineR » et « Facto extra ». Nous administrons un traitement CAH aux résultats issus de l'ACP et nous récupérons ainsi trois graphiques et une série de données. Le premier graphique est un dendrogramme qui est sectionné en quatre clusters (Figure 4). Le second graphique, en trois dimensions, est constitué du premier plan factoriel sur lequel sont montés le dendrogramme et les quatre clusters (Figure 5). Le troisième est le premier plan factoriel avec les clusters d'entreprise viticoles (Figure 6).

Le dendrogramme nous montre la distance entre les groupes. La première coupe se fait entre le cluster 4 et les trois autres clusters. Cette coupure crée un gain en inertie supérieure à 1,7 (Figure 4). Cette inertie est montrée par la hauteur de la branche du dendrogramme. Les deux sections supplémentaires qui créent les trois derniers groupes, créent des gains d'inertie de 0,7, 0,6 et de 0,5 (Figure 4). La Figure 6 illustre le premier plan factoriel avec les 4 clusters. Les groupes 1, 2 et 3 sont regroupés et nous allons procéder à l'interprétation.



Figure 6 : Premier plan factoriel avec 4 clusters.



Dans notre interprétation nous sommes aidés par les 50 entretiens qualitatifs que nous avons conduits auprès des entreprises viticoles sur leurs motivations d'arrachage. Nous avons tenté de retrouver au sein des 4 clusters ces 50 entreprises viticoles. Lors du nettoyage des données, nous avons éliminé les entrées avec des données manquantes. De ce fait 19 entreprises parmi les 50 n'apparaissent pas au sein d'un cluster. Nous avons créé le Tableau 3 qui décrit dans chaque cluster les motivations d'arrachages des entreprises pour lesquelles nous avons pu mettre en place un entretien qualitatif. À l'aide du Tableau 2 et Tableau 3 nous avons essayé de définir une typologie des entreprises arrachant des vignes.

– Grande entreprise en forte décroissance homogène selon l'âge : recherche d'opportunité ou valeur d'option.

Le groupe 4 a donc ses particularités. Et il est composé de 2 individus dont la taille est grande : 56.44 ha et 69.52 ha (Tableau 2). Ils arrachent 1/3 de leur superficie. Outre la distinction par la taille et leur pourcentage d'arrachage, ces exploitations ont la particularité d'arracher d'importantes quantités de vignes dans chaque catégorie d'âge. En moyenne, les deux entreprises arrachent 4 ha de vignes jeunes 0 à 15 ans, 6 ha de vignes de 16 à 30 ans, presque 10 hectares de vignes de 30 à 45 ans et 1,5 ha de vignes âgées de plus de 45 ans.

– Les sortants avec arrachage total ou quasi total de petites surfaces : les « vitexits ».

Le premier groupe est composé de 70 entreprises viticoles de petite taille allant de moins de 1 hectare à presque 6 hectares (Tableau 2). La particularité de ce groupe est son taux d'arrachage très élevé. Le taux d'arrachage médian est de 100 % et la moyenne de 82 % (plus de la moitié des entreprises arrachent la totalité de leur vignoble). Dans les arrachages, les jeunes vignes de moins de 15 ans sont quasi absentes. Les arrachages sont

principalement des vignes de 16-30 ans (en moyenne 0.5 hectares), suivi des vignes de 31-45 ans (en moyenne 0.4 hectares) et des vignes de plus de 45 ans (en moyenne 0.3 hectares).

Ce sont des viticulteurs probablement âgés sans successeurs ou sans repreneurs disposant de vignes âgées et de cépages traditionnels, ou peut-être des fractions d'exploitations scindées à l'occasion de l'arrachage subventionné pour être prioritaire dans l'obtention de la prime. Ils apparaissent dans toutes les catégories de motivation à l'arrachage (Cf. Tableau 3).

– *Désendettement et trésorerie : faible taux d'arrachage de vignes dans la force de l'âge.*

Le deuxième groupe est constitué de 122 entreprises viticoles, soit près des 2/3, d'une taille moyenne de 10 ha (Tableau 2). La taille des entreprises est assez hétérogène, car l'entreprise la plus petite est de 1,2 ha et la plus grande de 73 ha. Ce qui caractérise ce groupe est le pourcentage de vigne arrachée assez faible. La valeur médiane est de 13 % et la moyenne de 16 %. Les entreprises viticoles arrachent en proportion nettement moins que le premier groupe. En outre, elles arrachent principalement des vignes âgées de 16 à 30 ans (en moyenne 0,74 ha) et de 31 à 45 ans (en moyenne 0,48 ha). Pratiquement aucune vigne d'âge supérieur à 45 ans ou inférieur à 15 ans n'est arrachée. La valeur moyenne d'arrachage pour ces deux catégories est inférieure à 0,20 ha. Ce groupe n'est pas directement concerné par la politique d'élimination des entreprises obsolètes. Il tient à conserver une part significative de l'exploitation, quelle que soit sa taille. L'avantage majeur du recours à cette subvention est d'améliorer son endettement et sa trésorerie (Cf. Tableau 3). Il concerne des vigneronns endettés à la suite de la crise en 2004 avec notamment des emprunts pour le foncier.

– *Rajeunissement du capital productif : taux d'arrachage significatif de vieilles vignes.*

Le troisième groupe a la même hétérogénéité que le deuxième groupe et il est constitué de 19 entreprises d'une taille légèrement plus grande (Tableau 2). La taille moyenne est de 21,8 ha (min 3,9 ha et max 98 ha). La différence avec le groupe 2 se situe au niveau de la nature des vignes arrachées et de la proportion des vignes arrachées. C'est un groupe qui arrache en moyenne presque 1/3 de sa superficie (30,8 %) et principalement des vignes âgées de plus de 45 ans (en moyenne 2,8 ha). Les valeurs médianes des superficies arrachées dans les autres catégories d'âge de vigne sont de 0 ha.

**Tableau 2 : description des clusters « entreprise »**

	Cluster	0-15	16-30	31-45	46 +	ha arrachés	Sup total	%_arr	# entreprise	% d'effectif
Médiane	1	0	0.1905	0	0	1.0747	1.4650	1.0000	70	33 %
Moyenne		0.02499	0.5255	0.3973	0.3205	1.3552	1.7890	0.8212		
Médiane	2	0	0.3743	0	0	1.329	10.906	0.13239	122	57 %
Moyenne		0.1824	0.7389	0.4763	0.1469	1.671	15.075	0.16293		
Médiane	3	0	0	0	1.8860	2.860	18.217	0.23264	19	9 %
Moyenne		0.008421	0.6165	0.4997	2.3059	3.797	21.887	0.30849		
Médiane	4	4.096	6.284	6.284	1.6705	21.01	62.98	0.3303	2	1 %
Moyenne		4.096	6.284	9.667	1.6705	21.01	62.98	0.3303		

**Tableau 3 : distribution des entreprises par cluster et motivations d'arrachage**

Cluster	Départ du Fermier	Pas de repreneur	Rentabilité faible	Retraite	Santé	Trésorerie	Total
1	1	0	9	2	1	3	<b>16</b>
2	0	0	6	0	0	3	<b>9</b>
3	0	3	0	2	0	1	<b>6</b>
4	0	0	0	0	0	1	<b>1</b>
<b>Total</b>	<b>1</b>	<b>3</b>	<b>15</b>	<b>4</b>	<b>1</b>	<b>8</b>	

En première approximation on peut considérer que le groupe 1 rentre parfaitement dans le premier critère de l'objectif de la politique à savoir favoriser les départs à la retraite et que le groupe 2 répond plutôt au deuxième critère de la politique d'arrachage à savoir l'élimination de vieilles vignes et de cépages inadaptés. Les groupes 3 et 4 sont plus proches d'une opportunité d'amélioration de trésorerie.

## **4 Interprétation du second tableau issu d'une jointure par identité parcellaire et test du chi 2 sur les cépages.**

L'importation et le nettoyage du tableau de données de parcelles viticoles subissant l'arrachage de vignes ont été les premières étapes de notre travail. Une fois ces étapes accomplies, nous avons soumis nos données aux deux mêmes traitements statistiques ACP et CAH. Ces traitements nous ont fourni la série de graphiques et de tableau suivants.

### **4.1 Les variables et le premier plan factoriel.**

Au regard de la Figure 7, le cercle de corrélations des variables, nous voyons deux groupes de variables. Trois variables constituent un premier groupe : la superficie totale de l'exploitation viticole, les hectares de vignes arrachés, et la taille de la parcelle de vigne arrachée (respectivement : sup\_total, haarraches et ha\_vig\_arr.). Les flèches qui représentent les trois variables tendent vers le périmètre du cercle de rayon 1. Leur représentation graphique est donc correcte (Bialès and Fenneteau, 1998). Les trois variables sont corrélées, car elles pointent toutes dans la même direction. À elles trois, elles représentent quasiment la moitié de la variance (48,45 %) et constituent donc la première dimension du plan factoriel de l'ACP, comme le confirme le Tableau 4.

Le deuxième groupe de variable est constitué d'une seule variable, l'âge des vignes arrachées (age). Cette variable est représentée de façon excellente, car sa position touche quasiment le périmètre du cercle de rayon 1 (Bialès and Fenneteau, 1998). Elle n'est pas corrélée avec le premier groupe, car l'angle créé entre les deux groupes de valeur dans la figure 1 est de 90 degrés. Elle représente à seule 1/4 de la variance et elle contribue de façon quasi exclusive à l'établissement de la deuxième dimension (voir Tableau 4 : âge/dim2 = 0.98432441).

Nous pouvons donc conclure que le premier plan factoriel a en ordonnées l'âge des vignes arrachées et en abscisse une donnée mêlant la superficie de l'entreprise viticole, les

hectares de vignes arrachées et la taille des parcelles de vignes arrachées. Nous nous limiterons à l'analyse des deux premières dimensions du plan factoriel. La CHA nous aidera à préciser notre analyse

Figure 7 : Cercle des corrélations des variables

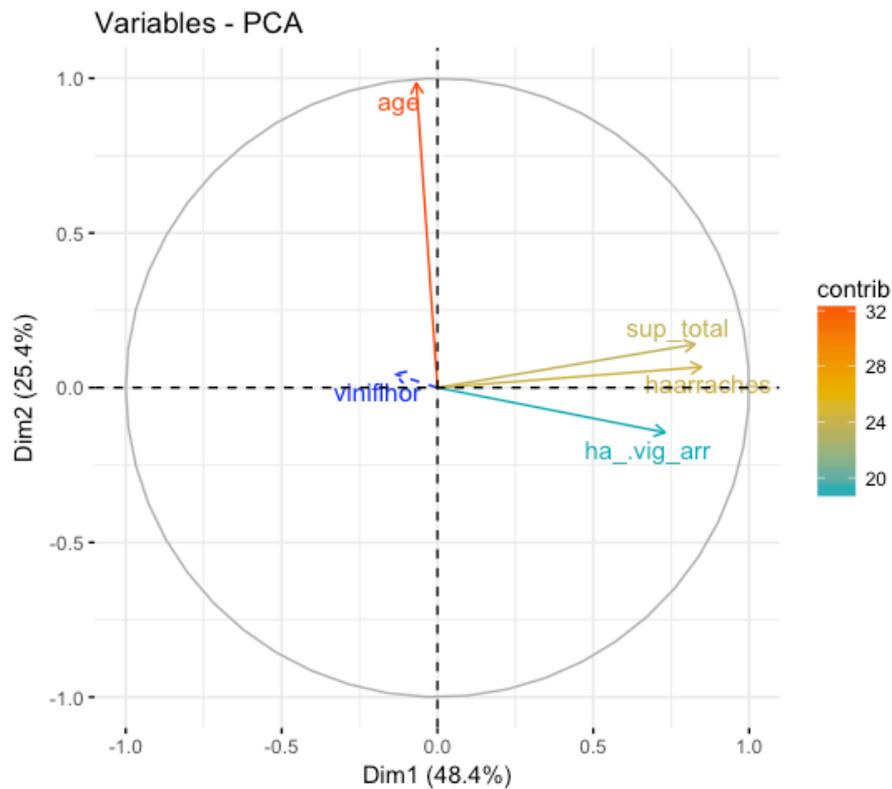


Tableau 4 : participation des variables aux dimensions factorielles

Variables	Dim.1	Dim.2	Dim.3	Dim.4
ha_vig_arr	0.73023304	-0.14479363	0.6655096	-0.05377255
age	-0.06780887	0.98432441	0.1613081	0.02207019
haarraches	0.8472229	0.06664245	-0.2371388	0.47067754
sup_total	0.82600024	0.3318759	-0.3318759	-0.43342087

#### 4.2 Individus, variables et classification hiérarchique ascendante.

Le traitement statistique par CAH est illustré par un dendrogramme reprenant l'ensemble des parcelles, visible sur la Figure 9. Ce graphique nous permet d'observer les gains d'inertie lors de la création de nouvelles classes. Chaque coupe du dendrogramme crée des groupes de plus en plus homogènes. En regardant le Figure 9, nous constatons que la première section crée le plus important gain d'inertie (1,4). Une coupe additionnelle de l'arbre ajoute

un gain de 0,5 en inertie. Trois coupes possibles sont encore exécutables et elles nous permettent de récupérer encore 0,3 d'inertie pour chacune d'entre elles (soit un total de 0,9). Les 5 coupes du dendrogramme donnent naissance à 6 clusters.

Ces 6 clusters sont représentés par 6 couleurs sur le premier plan factoriel sur la Figure 9. Une première analyse met en exergue deux groupes séparés sur l'axe 1 : le groupe 1 composé des clusters 1, 2, 3, 4 (noire, rouge, vert et bleu foncé) et le groupe 2 composé des clusters 5 et 6 (bleu ciel et rose). Ceci correspond à la première « coupe ». La différence entre ces deux groupes se définit par la taille de l'exploitation viticole et de la surface arrachée. Le groupe 1 réunit des entreprises viticoles plus petites et qui arrachent moins en valeur absolue en comparaison au groupe 2.

La deuxième coupe, sépare le groupe 1 en deux selon l'âge. Ainsi les clusters 2 et 4 (rouge et bleu) sont similaires et regroupent les parcelles dont les vignes arrachées sont plus jeunes. Les clusters 1 et 3 (verts et noirs) se ressemblent et représentent l'arrachage de vignes plus âgées. Les vignes du cluster 3 (vert) sont plus âgées que les vignes du cluster 1. Finalement, trois coupes supplémentaires divisent chacun de ces groupes et créent au total 6 groupes. L'illustration de ces 6 groupes est faite par le dendrogramme qui apparaît dans la figure 3 sur le premier plan factoriel. Il reconstitue les 6 clusters. La division entre les deux groupes de clusters est clairement décelable par la plus grande distance de « branches ».

Les clusters 1, 2 et 3 sont constitués de parcelles appartenant à des entreprises viticoles ayant approximativement la même taille, avec une taille médiane aux alentours de 7ha et qui arrachent en moyenne entre 25 % et 30 % de leur vignoble, soit des surfaces entre 1,5 et 2,5 ha. La caractéristique qui distingue ces trois groupes est essentiellement l'âge des cépages. La moyenne d'âge des vignes arrachées dans le cluster 2 est de 22 ans (rouge), dans le cluster 1 de 46 ans (noir) et dans le cluster 3 de 86 ans (vert).

Le cluster 4 (bleu) est composé d'entreprises viticoles de tailles légèrement plus grandes, avec une moyenne de 31 hectares et qui arrachent en moyenne 4,28 ha (13,8 %) de vignes d'âge médian et moyen d'approximativement 30 ans.

Les clusters 5 et 6 arrachent essentiellement des vignes âgées en moyenne entre 30 et 35 ans. La superficie de ces entreprises viticoles est nettement plus grande que celles constituant les autres groupes. Les entreprises viticoles du cluster 5 possèdent une superficie moyenne de 50 ha. L'arrachage moyen est de 22 %, le médian est de 10 %. Le cluster 6 est constitué que d'une seule entreprise viticole de 69,52 ha qui arrache 25 % de sa superficie.

Le Tableau 5 reprend l'essentiel de cette information et la Figure 10 représente les 6 clusters dans le premier plan factoriel (les couleurs des clusters ne sont pas les mêmes que dans la Figure 9)

Figure 8 : dendrogramme de 6 clusters « parcelles » issus de la CAH

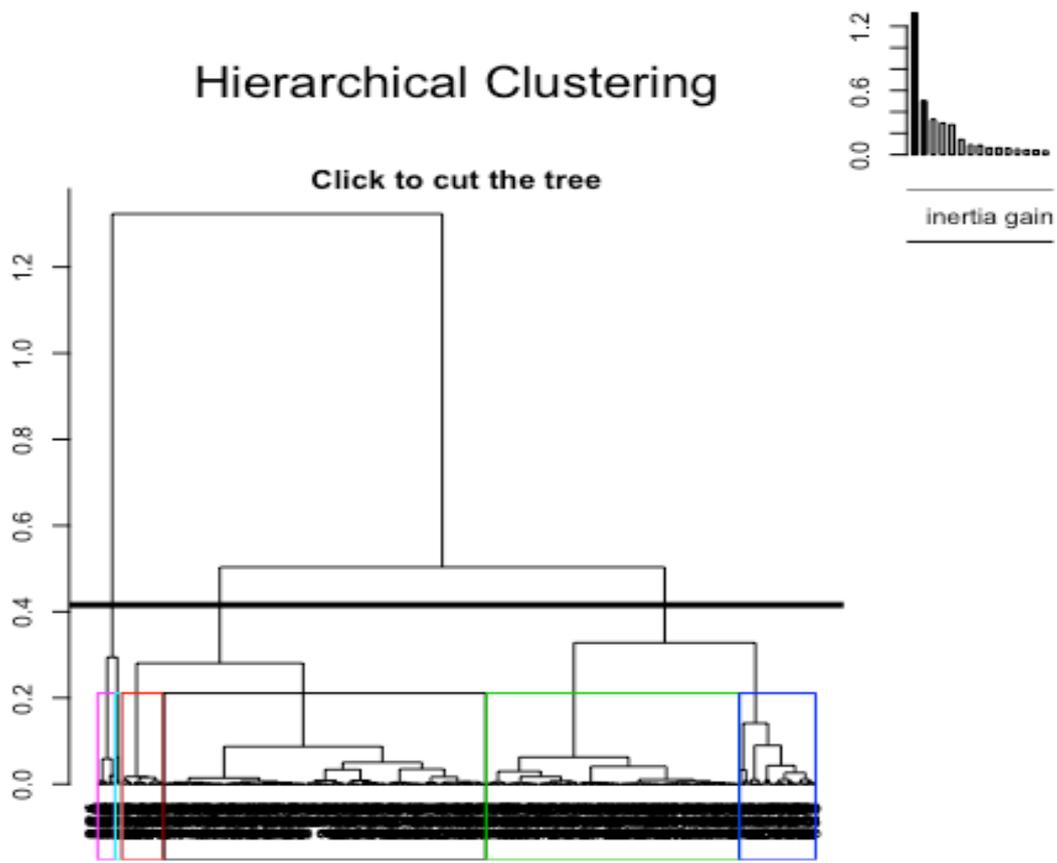


Figure 9 : 6 clusters « parcelles » sur le premier plan factoriel de l'ACP avec le dendrogramme de l'ACP

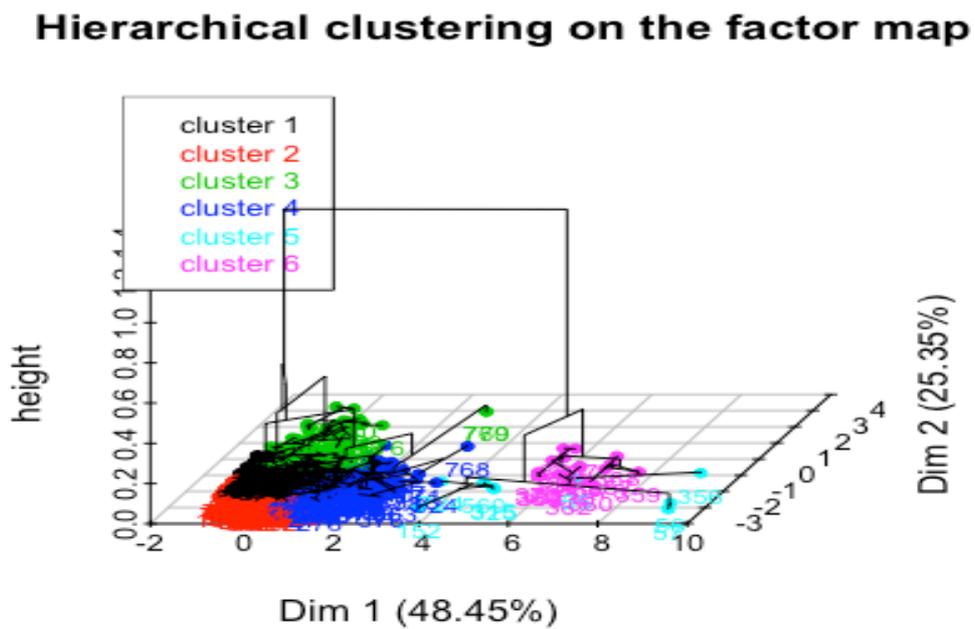


Figure 10 : 6 clusters « parcelles » issus de la CAH sur le premier plan factoriel de l'ACP

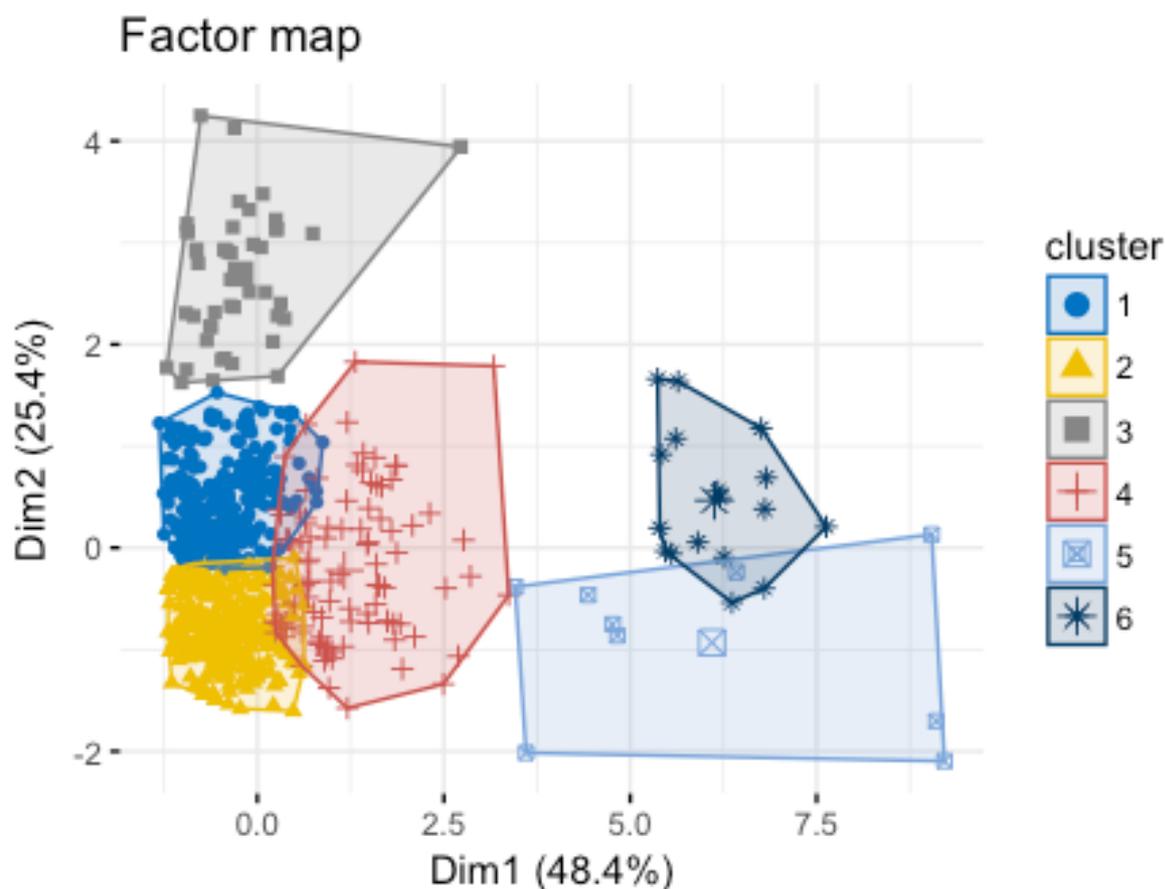


Tableau 5 : résumé des statistiques de 6 clusters par parcelles

Cluster	Âge moyen	Superficie arrachée moyenne ha	Superficie exploitation ha	% arraché	nombre de parcelle	nombre d'entreprise
Cluster 1	46,41	2,23	8,25	27,00 %	283	104
Cluster 2	22,96	2,04	6,77	30,13 %	311	119
Cluster 3	86,96	2,50	15,64	16,00 %	47	26
Cluster 4	30,76	4,28	31,00	13,80 %	104	42
Cluster 5	28,83	10,95	50,00	21,90 %	9	7
Cluster 6	35,00	25,16	69,52	36,20 %	17	1

### 4.3 Cépages et test du Chi2.

À l'issue de l'ACP et de la CAH, il nous reste encore une variable qualitative à analyser : les cépages. Dans notre premier article, nous avons une notion globale sur la distribution des cépages. Ici, nous pouvons appliquer un test du chi 2 à la variable cépage au sein de chaque cluster pour savoir s'il existe une surreprésentation ou une sous-représentation d'un cépage au sein d'un cluster.

Le cluster 1 est surreprésenté avec les cépages carignan, aramon, terret, carignan blanc et il est sous représenté avec les cépages mourvèdre, alicante, sauvignon blanc, grenache,

cinsault, merlot et syrah. En ce qui concerne le cluster 2, il existe une surreprésentation des cépages Merlot, syrah, cinsault, sauvignon blanc, grenache, et une sous-représentation des cépages terret, terret blanc, aramon et carignan. Au sein du cluster 3, le cépage aramon est surreprésenté et le cépage syrah est sous représenté. Cependant, ce dernier est surreprésenté dans le cluster 4. Finalement, aucun cépage n'a de particularité au niveau du test du chi 2 dans le cluster 5 et l'alicante est surreprésenté dans le cluster 6.

Les clusters 1, 3 et 6 regroupent des vignes arrachées de cépages traditionnels et d'un âge plutôt élevé. Les clusters 2 et 4 arrachent les cépages de dernières générations qualifiés d'améliorateurs.

## 5 Conclusion

Nous avons utilisé un outil assez sophistiqué qui permet d'obtenir une typologie des exploitations tenant compte de toutes les caractéristiques observées. En ce sens, cette typologie nous permet de dépasser l'observation des quatre groupes définis sur le graphique taux d'arrachage-surface utilisé dans notre premier article. Nous avons une meilleure estimation des effectifs des groupes répondant ou non aux deux objectifs de la politique d'arrachage. Le groupe 1 du cluster exploitation avec 70 entreprises représente 33 % de l'effectif et répond bien à l'objectif d'élimination des entreprises les moins performantes. Le groupe 3 du cluster « entreprise » avec 9 entreprises représentant 9 % de l'effectif répond bien au deuxième objectif d'amélioration du potentiel de production. Pour les deux autres groupes représentant plus de la moitié des entreprises on peut supposer qu'ils profitent du « windfall », en fait de l'opportunité des subventions pour refaire leur trésorerie, en arrachant plutôt une petite quantité de vignes jeunes et de cépage améliorateurs. Cette analyse ne fait pas apparaître de comportements totalement atypiques qui ne seraient pas apparus dans l'analyse descriptive initiale comme c'est souvent le cas en analyse de données. Par contre, les résultats diffèrent significativement au niveau des effectifs des groupes.

## 6 Bibliographie

- Batzeli, K., 2007. Rapport sur la réforme de l'organisation commune du marché du vin (2006/2109 (INI)). Parlement Européen, Strasbourg.
- Bialès, C., Fenneteau, H., 1998. Analyse statistique des données : Applications et cas pour le marketing. Ellipses Marketing, Paris.
- Centre d'Économie Rurale (CER) France, 2009. Dossier de presse : « La crise viticole clairement illustrée par cette étude » Synthèse de la 4ème édition de l'observatoire de la production viticole en Languedoc-Roussillon du 9 décembre 2009.
- Centre d'Économie Rurale (CER) Languedoc-Roussillon, 2006. Économie Languedoc-Roussillon : 90 % des exploitations viticoles seraient en difficulté financière. Vitisphère-Economie Dossiers 16/06/2006.
- Cornillon, P.-A., Husson, F., Jégou, N., Matzner-Lober, E., Collectif, 2012. Statistiques avec R,

- 3e édition revue et augmentée. ed. PU Rennes, Rennes.
- Delord, B., Montaigne, É., Coelho, A., 2015. Vine planting rights, farm size and economic performance: do economies of scale matter in the French viticulture sector? *Wine Economics and Policy* 4, 22–34.
- European Commission, 2006. Wine: Common Market Organisation, working paper. Brussels.
- Husson, F., Lê, S., Pagès, J., 2016. Analyse de données avec R, 2e édition revue et augmentée. ed. PU Rennes.
- Kassambara, M.A., 2017. Practical Guide To Principal Component Methods in R, 1 edition. ed. CreateSpace Independent Publishing Platform, United States.
- Lê, S., Josse, J., Husson, F., 2008. FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software* 25. <https://doi.org/10.18637/jss.v025.i01>
- Montaigne, É., Zadmehrán, S., Coelho, A., Messaoudène, Y., 2017. La politique d'arrachage a-t-elle répondu à ses objectifs en 2007 : Une étude de cas sur un échantillon départemental.