

# Déterminants biophysiques et historiques des AOC viticoles en Bourgogne

Jean-Sauveur AY                      Mohamed HILAL  
< [jean-sauveur.ay@inra.fr](mailto:jean-sauveur.ay@inra.fr) >      < [mohamed.hilal@inra.fr](mailto:mohamed.hilal@inra.fr) >

Unité Mixte de Recherche CESAER  
AgroSup / INRA / Univ. Bourgogne Franche-Comté  
26 boulevard du docteur Petitjean 21000 DIJON

*Data paper* version 0.2 du Mardi 04 juin 2019

## Résumé

Cet article présente la construction de données au niveau des parcelles cadastrales pour étudier statistiquement les liens entre les caractéristiques biophysiques (topographie, géologie, pédologie) et les appellations d'origine contrôlée (AOC). Sur les 31 communes de la Côte d'Or qui forment la Côte de Beaune et la Côte de Nuits, nous proposons une modélisation économétrique qui permet de classer l'ensemble des parcelles sur une échelle continue à partir de leurs caractéristiques biophysiques. Nous obtenons une persistance d'effets communaux que nous interprétons comme issus d'éléments historiques. Les données, méthodes et prédictions sont disponibles sous licence GNU GPL v3 sur <https://data.inra.fr/geoInd> et sont consultables par le biais d'une application sur <http://github.com/jsay/geoInd/>.

**Mots-clés:** Économie viti-vinicole ; signes de qualité ; recherche reproductible ; système d'information géographique ; modélisation économétrique.

## Table des Matières

<b>1</b>	<b>Introduction</b>	<b>2</b>	<b>4</b>	<b>Modèle économétrique</b>	<b>14</b>
<b>2</b>	<b>Construction des données</b>	<b>3</b>	4.1	Estimation du modèle	14
2.1	Les AOC au niveau des parcelles	3	4.2	Effets des variables biophysiques	15
2.2	Enrichissement de la topographie	4	4.3	Effets communaux	15
2.3	Enrichissement de la géologie	5	4.4	Prédiction du score et classifications	16
2.4	Enrichissement de la pédologie	6	4.5	Aggrégation	19
2.5	Enrichissement des AOC de 1936	7	<b>5</b>	<b>Application cartographique</b>	<b>19</b>
2.6	Enrichissement des lieux dits	8	5.1	R code	19
2.7	Enregistrement de la base	9	5.2	Global	21
<b>3</b>	<b>Statistiques descriptives</b>	<b>11</b>	5.3	Server	22
3.1	Filtrage des données	11	5.4	ui	22
3.2	Distribution des AOC	11	<b>6</b>	<b>Conclusion</b>	<b>23</b>
3.3	Les AOC historiques	12	<b>7</b>	<b>Bibliographie</b>	<b>23</b>
3.4	Autre graphique	12	<b>A</b>	<b>Annexes</b>	<b>24</b>
3.5	yop	14			

# 1 Introduction

Les appellations d'origine contrôlée (AOC) viticoles en Bourgogne résultent de processus historiques complexes au cours desquels les parcelles de vigne ont été classifiées selon leurs caractéristiques biophysiques et les rapports économiques, politiques et sociaux en vigueur (Garcia, 2011; Wolikow and Jacquet, 2011). La classification actuelle est issue de plusieurs siècles de culture de la vigne, de production de vin et de négociation sur les dénominations. Ces trois ensembles de pratiques forment les usages loyaux et constants définis dans la doctrine de l'institut national de l'origine et de la qualité (INAO) qui est elle-même un processus historique (Capus, 1947; Humbert, 2011). La complexité des informations contenues dans la référence au lieu de production et leur évolution dans le temps sont à la fois la force et la faiblesse des AOC, car elles permettent de simplifier les nombreux déterminants de la qualité des vins au risque d'une faible pertinence et d'une opacité croissante pour les échanges de vin.

La question de la transmission de l'information sur la qualité des biens fait l'objet d'une littérature économique abondante (Coestier and Marette, 2004). L'asymétrie d'information y est décrite comme une défaillance de marché qui invalide le premier théorème du bien-être et donc diminue l'efficacité de l'allocation par le libre marché. De plus, le recours aux indications géographiques apparaît typiquement comme une solution partielle qui segmente artificiellement la production et génère des rentes non justifiées pour les producteurs au détriment des consommateurs. Nous nous concentrons dans cet article sur la capacité des AOC à simplifier l'information sur les caractéristiques des lieux de production, en laissant de côté la question de la pertinence de cette information pour le marché. Nous proposons de démêler statistiquement les déterminants biophysiques et historiques des AOC par l'utilisation de données exhaustives à l'échelle des parcelles du cadastre. Nous proposons également d'affiner l'information contenue dans les AOC actuelles par une classification continue des vignes corrigée économétriquement des effets communaux issus de l'histoire. Ces méthodes et résultats sont librement utilisables par le biais d'une application cartographique.

Le travail sur les données consiste à apparier les informations biophysiques des parcelles cadastrales aux AOC par l'utilisation d'un système géographique d'information. La parcelle cadastrale est l'unité géographique de base qui permet le calcul de variables topographiques (IGN 5 m), de variables géologiques (BRGM) et de variables pédologiques (RPB). Cette procédure de construction des données est décrite en détails dans la Section XX avec les codes utilisés. Les données se limitent actuellement aux 31 communes carte de la zone pas les hautes cotes ni le châtillonnais comme dans la Figure XX. quelques statistiques descriptives sont présentées dans la section XX.

La section XX présente le détails de l'estimation du modèle économétrique présenté plus extensivement dans un article associé (Ay, 2019). Il s'agit d'un modèle ordonné additif généralisé (OGAM) qui a recours à une variable latente. Nous présentons également une application de cette base de données dans la section XX.

Ce document contient le code R, packages, github, etc. Les bases de données sources qui entrent dans le travail sont disponibles auprès des auteurs sur demande. Reprendre les liens vers le data verse de l'INRA, les base brutes peuvent être obtenues sur demande auprès des auteurs. La construction des données est intéressante pour les procédures et la description des variables.

## 2 Construction des données

### 2.1 Les AOC au niveau des parcelles

L'unité géographique de base est la parcelle cadastrale dont la géométrie est issue de la BD parcellaire de l'IGN version X.XX téléchargée le XX/XX/2018 à l'adresse [XX](#). Ces données sont sous licence XX. Trois traitements ont été effectués au préalable et ne sont pas reportés en détail ici. Nous avons calculé avec un système d'information géographique les caractéristiques géométriques (surface, périmètre, et distance maximale entre deux sommets). Nous avons créé un identifiant pour apparier les parcelles avec les données du modèle numérique de terrain XX. Nous avons apparier les délimitations parcellaire des AOC Viticoles de l'INAO disponible à l'adresse <https://www.data.gouv.fr/fr/datasets/delimitation-parcellaire-des-aoc-viticoles-de-linao> sous licence ouverte. Le résultat est disponible sous format shapefile sur le serveur data de l'INRA <https://data.inra.fr/geoInd/GeoCad.zip> sous licence GNU GPL v3. Le déplacement de ces fichiers géographiques dans un répertoire Carto/ à la racine permet sa lecture dans R comme présenté ci-dessous.

---

```
library(sp) ; library(rgdal)
Geo.Cad <- readOGR("./Carto", "GeoCad")
sapply(Geo.Cad@data, function(x) sum(is.na(x)))
```

---

```
OGR data source with driver: ESRI Shapefile
Source: "/home/jsay/geoInd/Carto", layer: "GeoCad"
with 110350 features
It has 16 fields
```

IDU	CODECOM	AREA	PERIM	MAXDIST	PAR2RAS	PAOC	BGOR
0	0	0	0	0	0	49718	49718
BOUR	VILL	COMM	PCRUR	GCRUR	AOC	AOCtp	AOC1b
49718	49718	49718	49718	49718	0	49718	49718

Le fichier contient 110 350 parcelles et 16 variables que la Table 1 suivante présente plus en détails. Les 49 718 valeurs omises qui apparaissent en sortie correspondent aux parcelles hors AOC, elles sont codées 0 dans la hiérarchie représentée par la variable AOC. L'information brute issue de la superposition avec la couche INAO est présente dans les variables PAOC à GCRUR. Nous l'avons résumé dans les trois variables AOC, AOCtp et AOC1b qui sont plus opérationnelles. Selon le principe des replis de la doctrine de l'INAO, les parcelles d'un niveau hiérarchique supérieur peuvent toujours être revendiquées dans un niveau inférieur. Cela produit la présence de plusieurs AOC sur une même parcelle selon les variables issues de la superposition des couches alors qu'il est interdit de revendiquer des AOC différentes. La variable AOC représente l'AOC maximale à laquelle la parcelle peut prétendre qui représente très majoritairement l'AOC qui est effectivement revendiquée. Par contre, les informations présentes sur l'étiquette des vins peuvent être des appellations ou des dénominations au sein du système des AOC (même si cette distinction n'est pas toujours claires pour les individus, nous utilisons AOC comme le terme générique qui englobe les deux en précisant lorsque c'est nécessaire). Le libellé AOC1b renvoi généralement l'appellation sauf pour les "Bourgognes Régionaux" (ou la dénomination "Bourgogne Côte d'Or" est prédominante) et les "Premiers Crus" (qui ont chacun une dénomination qui permet de les distinguer). La commande `table(Geo.Cad$AOC, Geo.Cad$AOCtp)` permet de rendre compte de cette structuration des variables.

Table 1: **Nom, type et description des variables disponibles au niveau parcellaire**

NOM	TYPE	DESCRIPTION
IDU	<i>Caractère</i>	Identifiant cadastral de la parcelle (14 caractères)
CODECOM	<i>Caractère</i>	Code INSEE de la commune d'appartenance (5 caractères)
AREA	<i>Numérique</i>	Surface calculée de la parcelle (en mètres carrés)
PERIM	<i>Numérique</i>	Périmètre calculé de la parcelle (en mètres)
MAXDIST	<i>Numérique</i>	Distance maximale calculée entre deux sommets (en mètres)
PAR2RAS	<i>Numérique</i>	Identifiant pour appariement avec le modèle numérique de terrain
PAOC	<i>Indicatrice</i>	1 si la parcelle est dans au moins une AOC
BGOR	<i>Indicatrice</i>	1 si la parcelle est dans le niveau Coteaux Bourguignon
BOUR	<i>Indicatrice</i>	1 si la parcelle est dans le niveau Bourgogne Régional
VILL	<i>Indicatrice</i>	1 si la parcelle est dans le niveau Bourgogne Village
COMM	<i>Indicatrice</i>	1 si la parcelle est dans le niveau Bourgogne Communal
PCRU	<i>Indicatrice</i>	1 si la parcelle est dans le niveau Premier Cru
GCRU	<i>Indicatrice</i>	1 si la parcelle est dans le niveau Grand Cru
AOC	<i>Numérique</i>	Rang de la parcelle dans la hiérarchie des AOC (entre 0 et 5)
AOCtp	<i>Caractère</i>	Appel si le libellé est une appellation, Denom pour dénomination
AOC1b	<i>Caractère</i>	Libellé de l'appellation ou de la dénomination selon la variable AOCtp

## 2.2 Enrichissement de la topographie

Les données issues du modèle numérique de terrain sont disponibles au format raster sur le serveur data de l'INRA <https://data.inra.fr/geoInd/vitidem.csv> sous licence GNU GPL v3. Un premier traitement non reporté a été l'attribution de l'identifiant PAR2RAS aux cellules une par superposition avec la géographie du parcellaire. Nous avons ensuite enrichi les données raster d'un mode d'occupation des sol (SOURCE) et d'une perméabilité calculée (SOURCE). Nous avons enfin calculé les variables topographiques que sont l'altitude, la pente, l'exposition et les radiations solaires (détails en Annexe). À partir des plus de 14 millions de cellules pour 13 variables, le code ci-dessous calcul les moyennes à l'échelle des parcelles.

```
library(data.table)
dim(Dat.Ras <- fread("./Data/DatRas.csv"))
Cad.Ras <- Dat.Ras[, lapply(.SD, mean),
                     by= list(PAR2RAS), .SDcols= names(Dat.Ras)[ -c(1, 4)]]
Geo.Ras <- merge(Geo.Cad, Cad.Ras, by= "PAR2RAS")
sapply(Geo.Ras@data[, 17: 26], function(x) sum(is.na(x))); rm(Dat.Ras)
```

data.table 1.11.4 Latest news: <http://r-datatable.com>

```
[1] 14253070      13
```

XL93	YL93	NOMOS	URBAN	FOREST	WATER	DEM	SLOPE	ASPECT	SOLAR
2096	2096	2096	2096	2096	2096	2096	2096	2096	2096

Le détails des variables issue du fichier raster est disponible dans la Table 2 ci-dessous. Nous obtenons 2 096 parcelles pour lesquelles le code PAR2RAS ne s'apparie à aucune cellule raster. Ces parcelles sont de

très petites parcelles avec des géométrie particulières et font penser à des "erreurs" du cadastre. Nous les enlèverons de l'analyse sachant que cela revient à enlever 2.7 ha, moins de 0.01 % de la surface totale.

Table 2: **Nom, type et description des variables issues des données raster**

NOM	TYPE	DESCRIPTION
XL93	<i>Numérique</i>	Latitude du centroïde de la parcelle (système Lambert 93)
YL93	<i>Numérique</i>	Longitude du centroïde de la parcelle (système Lambert 93)
NOMOS	<i>Numérique</i>	Part de la parcelle hors du mode d'occupation des sol (entre 0 et 1)
URBAN	<i>Numérique</i>	Part de la parcelle en usage urbain (entre 0 et 1)
FOREST	<i>Numérique</i>	Part de la parcelle en usage forestier (entre 0 et 1)
WATER	<i>Numérique</i>	Part de la parcelle en eau (entre 0 et 1)
DEM	<i>Numérique</i>	Altitude moyenne de la parcelle (en mètres)
SLOPE	<i>Numérique</i>	Pente moyenne de la parcelle (en degrés)
ASPECT	<i>Numérique</i>	Exposition moyenne de la parcelle (en degrés)
SOLAR	<i>Numérique</i>	Radiation solaire moyenne sur la parcelle (en Joules)
PERMEA	<i>Numérique</i>	Perméabilité des sols moyenne (entre 0 et 4)

## 2.3 Enrichissement de la géologie

Les données géologiques sont issues de la Bd Charm-50 du BRGM à l'échelle 1/50 000 disponible sur le site <http://infoterre.brgm.fr> sous licence Ouverte. Nous utilisons ici une extraction du fichier GEO050K\_HARM\_021\_S\_FGEOL\_CGH\_2154 effectuée en avril 2019 pour le département de la Côte d'Or. Le seul travail non reporté est une sélection des nombreuses variables disponibles sur la base des valeurs omises et sur la redondance de l'information qu'elle contiennent. Le résultat est disponible sous format shapefile sur le serveur data de l'INRA à l'adresse <https://data.inra.fr/geoInd/GeolMap.zip> sous licence GNU GPL v3.

```

Geol.Map <- readOGR("./Carto/", "GeolMap")
Pts.Cad <- SpatialPoints(Geo.Ras, proj4string= CRS(proj4string(Geol.Map)))
Geo.Ras@data <- cbind(Geo.Ras@data, over(Pts.Cad, Geol.Map))
sapply(Geo.Ras@data[, 28: 43], function(x) sum(is.na(x)))

```

OGR data source with driver: ESRI Shapefile  
Source: "/home/jsay/geoInd/Carto", layer: "GeolMap"  
with 13960 features  
It has 16 fields

CODE	NOTATION	DESCR	TYPEGEOL	APLOCALE	TYPEAP
31	31	31	31	862	862
GEOLNAT	ISOPIQUE	AGEDEB	ERADEB	SYSDEB	LITHOLOGIE
31	31	31	31	31	31
DURETE	ENVIRONMT	GEOCHIMIE	LITHOCOM		
69	31	31	69		

Le code précédent permet d'apparier les 13 960 polygones utilisés pour décrire la géologie avec les

parcelles cadastrales à partir de leurs centroïdes. La faible taille moyenne des parcelles viticoles (moins de 0.2 ha de moyenne) permet de s'assurer de la validité de cette procédure. Les détails des 16 variables géologiques sont disponibles dans la Table 3 suivante, bien que certaines informations manquent encore (une demande est en cours auprès du BRGM). Le manque de détails n'est pas fondamental pour l'analyse économétrique car les variables géologiques seront intégrées par des effets fixes qui permettent de contrôler l'hétérogénéité sans entrer dans les détails des effets. Les parcelles non appariées sont peut nombreuses (entre 31 et 862 selon les variables) et seront négligées dans l'analyse économétrique sans conséquence.

Table 3: **Nom, type et description des variables issues des données géologiques**

NOM	TYPE	DESCRIPTION
CODE	<i>Caractère</i>	Code de la géologie (31 modalités)
NOTATION	<i>Caractère</i>	Notation géologie (31 modalités)
DESCR	<i>Caractère</i>	Description géologie (31 modalités)
TYPEGEOL	<i>Caractère</i>	Type superficiel (4 modalités)
APLOCALE	<i>Caractère</i>	Colluvions, Eboulis, etc. (28 modalités)
TYPEAP	<i>Caractère</i>	Type de formation (7 modalités)
GEOLNAT	<i>Caractère</i>	Nature Géologique (3 modalités)
ISOPIQUE	<i>Caractère</i>	Faciès des couches (4 modalités)
AGEDEB	<i>Caractère</i>	Age de la couche (24 modalités)
ERADEB	<i>Caractère</i>	Céno ou Mésol (2 modalités)
SYSDEB	<i>Caractère</i>	Age autre (5 modalités)
LITHOLOGIE	<i>Caractère</i>	Litho (16 modalités)
DURETE	<i>Caractère</i>	Dureté (3 modalités)
ENVIRONMT	<i>Caractère</i>	Environnement (9 modalités)
GEOCHIMIE	<i>Caractère</i>	Géochimie (5 modalités)
LITHOCOM	<i>Caractère</i>	Litho détaillée (30 modalités)

## 2.4 Enrichissement de la pédologie

La couche pédologique est extraite du Référentiel Pédologique de Bourgogne : Régions naturelles, pédo-paysage et sols de Côte d'Or (étude 25021) à l'échelle 1/250 000, compatible avec la base de données nationale DoneSol. La localisation des types de sol présents sur la zone s'opère par les Unités Cartographiques de Sols, appelées aussi Pédo-paysages qui sont des polygones plutôt homogènes mais qui contiennent différents types de sols sans que ces derniers puissent être localisés plus précisément (Ay, 2011). En l'absence de données plus fines spatialement, les données parcellaires seront enrichies des codes des unités cartographiques que l'on enrichit avec des variables de description des sols. Les intitulés des UCS sont obtenus par un travail manuel reporté en Annexe (par le site <https://bourgogne.websol.fr/carto>).

```
Pedo.Map <- readOGR("./Carto", "PedoMap")
Geo.Ras@data <- cbind(Geo.Ras@data, over(Pts.Cad, Pedo.Map))
Geo.Ras@data[, c(45: 48, 50: 55)] <-
  apply(Geo.Ras@data[, c(45: 48, 50: 55)], 2, as.numeric)
sapply(Geo.Ras@data[, 44: 56], function(x) sum(is.na(x)))
```

OGR data source with driver: ESRI Shapefile

Source: "/home/jsay/geoInd/Carto", layer: "PedoMap"  
 with 194 features  
 It has 13 fields

NOUC	SURFUC	TARG	TSAB	TLIM	TEXTAG	EPAIS	TEG	TMO	RUE
14645	14645	14645	14645	14645	14645	14645	14645	14645	14645
RUD	OCCUP	DESCRp							
14645	14645	14645							

Il apparaît que les descriptions des Pédopaysages combinent des caractéristiques topographiques (Plaines, massifs, piedmonts), des caractéristiques d'occupation (forestiers, vignoble) et des caractéristiques géologiques (plio-pléistocènes, calcaires). Elles sont décrites en détails dans la Table 4 ci-dessous. Le redondance de ce découpage avec les variables topographiques, le découpage géologique et le mode d'occupation des sols se pose effectivement. Les valeurs manquantes correspondent *a priori* aux espaces urbanisés (mais cela ne se retrouve pas vraiment à partir du MOS).

Table 4: Nom, type et description des variables issues des données pédologiques

NOM	TYPE	DESCRIPTION
NOUC	<i>Caractère</i>	Numéro de l'unité cartographique (2 caractères)
SURFUC	<i>Numérique</i>	Surface de l'unité cartographique (en hectares)
TARG	<i>Numérique</i>	Taux d'argile de l'unité typologique dominante (pourcentage)
TSAB	<i>Numérique</i>	Taux de sable de l'unité typologique dominante (pourcentage)
TLIM	<i>Numérique</i>	Taux de limons de l'unité typologique dominante (pourcentage)
TEXTAG	<i>Caractère</i>	Classes de textures agrégées en 9 modalités (voir Ay, 2011)
EPAIS	<i>Numérique</i>	Épaisseur des sols de l'unité typologique dominante (centimètre)
TEG	<i>Numérique</i>	Taux d'éléments grossiers de l'unité typologique dominante (pour mille)
TMO	<i>Numérique</i>	Taux de Matière organique de l'unité typologique dominante (pourcentage)
RUE	<i>Numérique</i>	Réserve Utile par excès de l'unité typologique dominante (millimètre)
RUD	<i>Numérique</i>	Réserve Utile par défaut de l'unité typologique dominante (millimètre)
OCCUP	<i>Numérique</i>	Part de l'unité typologique dominante dans l'unité carto (entre 0 et 1)
DESCRp	<i>Caractère</i>	Libellé de la classe pédologique en 33 modalités

## 2.5 Enrichissement des AOC de 1936

Les AOC en vigueur en 1936 à la création de l'INAO ont été obtenues de la Maison des Sciences de l'Homme de Dijon (Licence?? avec l'aide de Florian Humbert). Un travail préalable a été effectué sur les AOC de 1936 afin de compiler les différentes années de 1936 à 1940. La localisation est effectuée par le centroïde des parcelles cadastrales car la géométrie des polygones ne correspond pas parfaitement (à la fois par la numérisation et parce que le cadastre a changé). Encore une fois, la faible taille des parcelles permet d'avoir confiance en cette procédure d'appariement.

```
Hist.Aoc <- readOGR("Carto/", "Aoc1936")
Geo.Ras@data <- cbind(Geo.Ras@data, over(Pts.Cad, Hist.Aoc))
sapply(Geo.Ras@data[, 57: 58], function(x) sum(is.na(x)))
```



```
OGR data source with driver: ESRI Shapefile
Source: "/home/jsay/geoInd/Carto", layer: "Aoc1936"
with 56 features
It has 2 fields
AOC36lab AOC36lv1
      70      70
```

Nous obtenons des aires sensiblement plus réduites que les actuelles, 27% au lieu de 55% trouvés ci-dessus. Hormis le creux de 1938, entre 10 et 15% des parcelles sont classées chaque années, sachant qu'il y a du double compte. Les premiers crus n'apparaissent pas car ils n'existaient pas à l'époque (création en 1948). le décret instaurant les Premiers Crus ne fut toutefois adopté qu'en 1943. Deux classements historiques servirent de principales références à la désignation de ces ceux-ci: celui de Jules Laval de 1855 et le Classement du Comité d'Agriculture et de Viticulture de l'Arrondissement de Beaune de 1860.

Table 5: **Nom, type et description des variables issues des AOC de 1936**

NOM	TYPE	DESCRIPTION
AOC36lab	<i>Caractère</i>	Libellé de l'appellation en 1936 (56 modalités)
AOC36lv1	<i>Caractère</i>	Rang de la parcelle dans la hiérarchie des AOC (entre 0 et 5)

## 2.6 Enrichissement des lieux dits

Il s'agit ici d'inclure de l'information cadastrale à partir des sources [data.gouv.fr](http://data.gouv.fr). Nous utilisons le Plan Cadastral Informatisé Vecteur (Format EDIGÉO, <https://cadastre.data.gouv.fr/datasets/plan-cadastral-informatise>) téléchargé pour la Côte d'Or (21) le <2019-01-13 dim.>. License ouverte Etalab. La difficulté avec les lieux dit est qu'ils doivent être croisés avec les communes car un même nom lieu dit peut être présent sur plusieurs communes. Comme la géométrie des lieux dits et des parcelles colle parfaitement, nous pouvons enrichir les données parcellaires directement par le centroïde. Ajout <2019-01-23 mer.>, des données communales, nous extrayons également les coordonnées des chefs-lieux pour calculer une distance à vol d'oiseaux, la population (peuvent être des sur-identifications sur le land use) et la distinction Côte de Beaune / Côtes de Nuits. Nous enregistrons également une shapefile MapCom qui permet de cartographier les contours communaux dans les figures.

```
Lieu.Dit <- readOGR("./Carto/", "LieuDit")
Geo.Ras@data <- cbind(Geo.Ras@data, over(Pts.Cad, Lieu.Dit[, -1]))
sapply(Geo.Ras@data[, 59: 68], function(x) sum(is.na(x)))
```

```
OGR data source with driver: ESRI Shapefile
Source: "/home/jsay/geoInd/Carto", layer: "LieuDit"
with 3285 features
It has 11 fields
```

```
LIEUDIT  CLDVIN  LIBCOM  XCHF  YCHF  ALTCOM  SUPCOM
      4494      4494      4494      4494      4494      4494      4494
POPCOM CODECANT  REGION
      4494      4494      4494
```



Pour 4% des parcelles, aucun lieu dit n'a été apparié. Ces parcelles se concentrent sur les communes de Chenôve, Marsannay-la-Côte et Beaune (Corgoloin dans une moindre mesure). Ces "trous" apparaissent déjà dans le fichier source et ne sont donc pas un résultat de l'appariement. Ils semblent être des espaces bâtis sur la carte, mais ce n'est toujours pas confirmé par le MOS.

Table 6: **Nom, type et description des variables issues des lieux dits**

NOM	TYPE	DESCRIPTION
LIEUDIT	<i>Caractère</i>	Libellé du lieu dit de la parcelle (2691 modalités)
CLDVIN	<i>Caractère</i>	Identifiant du lieu dit de la parcelle (2691 modalités)
LIBCOM	<i>Caractère</i>	Libellé de la commune de la parcelle (31 modalités)
XCHF	<i>Numérique</i>	Latitude du chef-lieu de la commune (système Lambert 93)
YCHF	<i>Numérique</i>	Longitude du chef-lieu de la commune (système Lambert 93)
ALTCOM	<i>Numérique</i>	Altitude du point culminant de la commune (mètre)
SUPCOM	<i>Caractère</i>	Superficie de la commune de la parcelle (hectare)
POPCOM	<i>Numérique</i>	Population de la commune de la parcelle en 2015 (millier d'hab)
CODECANT	<i>Caractère</i>	Identifiant du canton d'appartenance (2 caractères)
REGION	<i>Caractère</i>	Region viticole (CDB Côte de Beaune, CDN Côte de Nuits)

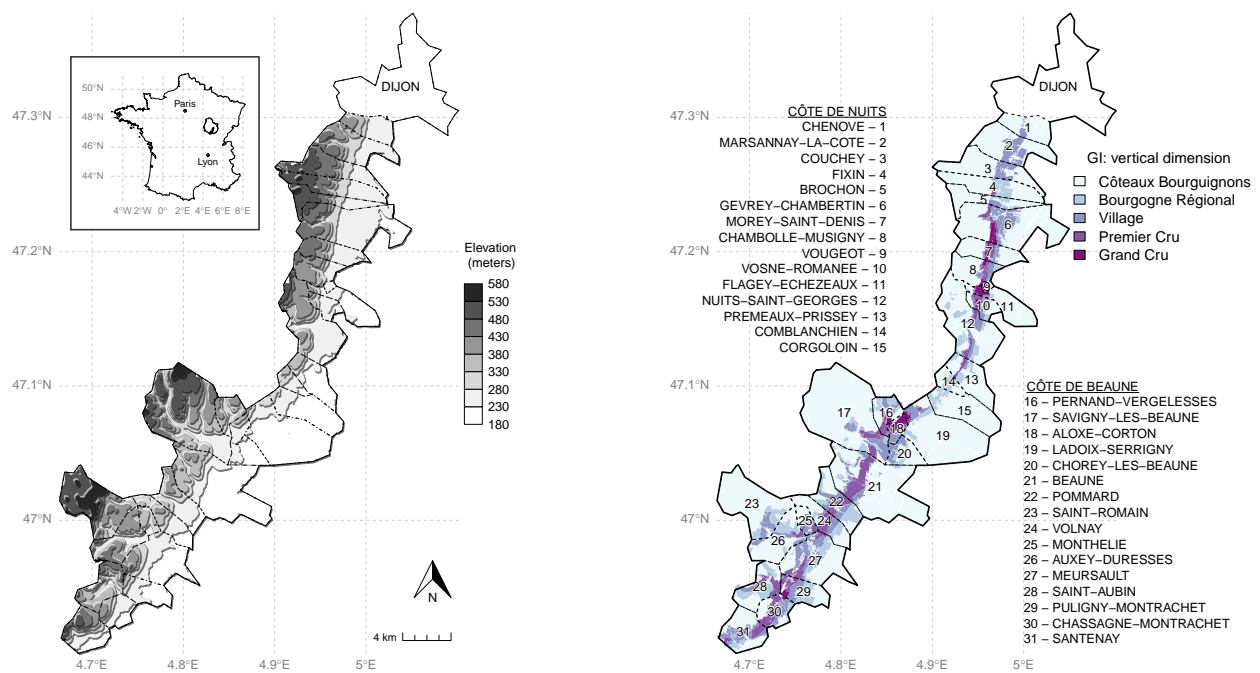
## 2.7 Enregistrement de la base

Pour l'instant, on est à moins de 500 Mo. correspond à l'unesco? <https://whc.unesco.org/fr/list/1425/>

```
dim(Geo.Ras)
save(Geo.Ras, file= "Inter/GeoRas.Rda")
writeOGR(Geo.Ras, "Carto/", "GeoRas", driver= "ESRI Shapefile")
```

[1] 110350 68

Figure 1: Vignobles de la Côte d'Or, topographie et appellations d'origine contrôlées  
blabla.



## 3 Statistiques descriptives

### 3.1 Filtrage des données

Parmi les fichiers XX disponible sur le serveur data de l'INRA, nous partons du fichier `GeoRas.Rda` que l'utilisateur doit placer dans un répertoire `Inter/` à la racine pour pouvoir utiliser le logiciel R (cite). La première procédure à exécuter est présentée ci-dessous. Elle consiste à:

- Recoder les codes communaux selon le gradient Nord-Sud
- Calculer la distance au chef lieu de la commune
- Centrer réduire la variable sur les rayonnements solaires
- Recoder la variable exposition en catégories
- Re-projeter les coordonnées dans le système WGS84
- Enlever les valeurs manquantes de la base de données

---

```
load("Inter/GeoRas.Rda")
tmp <- unique(Geo.Ras$LIBCOM[order(Geo.Ras$YCHF, decreasing= TRUE)])
Geo.Ras$LIBCOM <- factor(Geo.Ras$LIBCOM, levels= tmp)
Geo.Ras$DISTCHF <- sqrt((Geo.Ras$XL93- Geo.Ras$XCHF* 100)^2
+ (Geo.Ras$YL93- Geo.Ras$YCHF* 100)^2)
Geo.Ras$RAYAT <- (Geo.Ras$SOLAR- mean(Geo.Ras$SOLAR, na.rm= TRUE))/
sd(Geo.Ras$SOLAR, na.rm= TRUE)
Geo.Ras$EXPO <- cut(Geo.Ras$ASPECT,
breaks= c(-2, 45, 90, 135, 180, 225, 270, 315, 360))
GR84 <- spTransform(Geo.Ras, CRS("+proj=longlat +ellps=WGS84"))
dd <- coordinates(GR84) ; Geo.Ras$X= dd[, 1] ; Geo.Ras$Y= dd[, 2]
dim(Reg.Ras <- subset(Geo.Ras, !is.na(AOC1b) & !is.na(DEM) & !is.na(DESCR)
& !is.na(RUD) & !is.na(AOC36lab) & !is.na(REGION)))
```

---

```
[1] 59113      73
```

La limitation aux parcelle ayant des AOC contribue largement à la diminution du nombre d'information (perte de XX contre XX pour les valeurs manquantes).

### 3.2 Distribution des AOC

Les deux dimensions des indications géographiques.

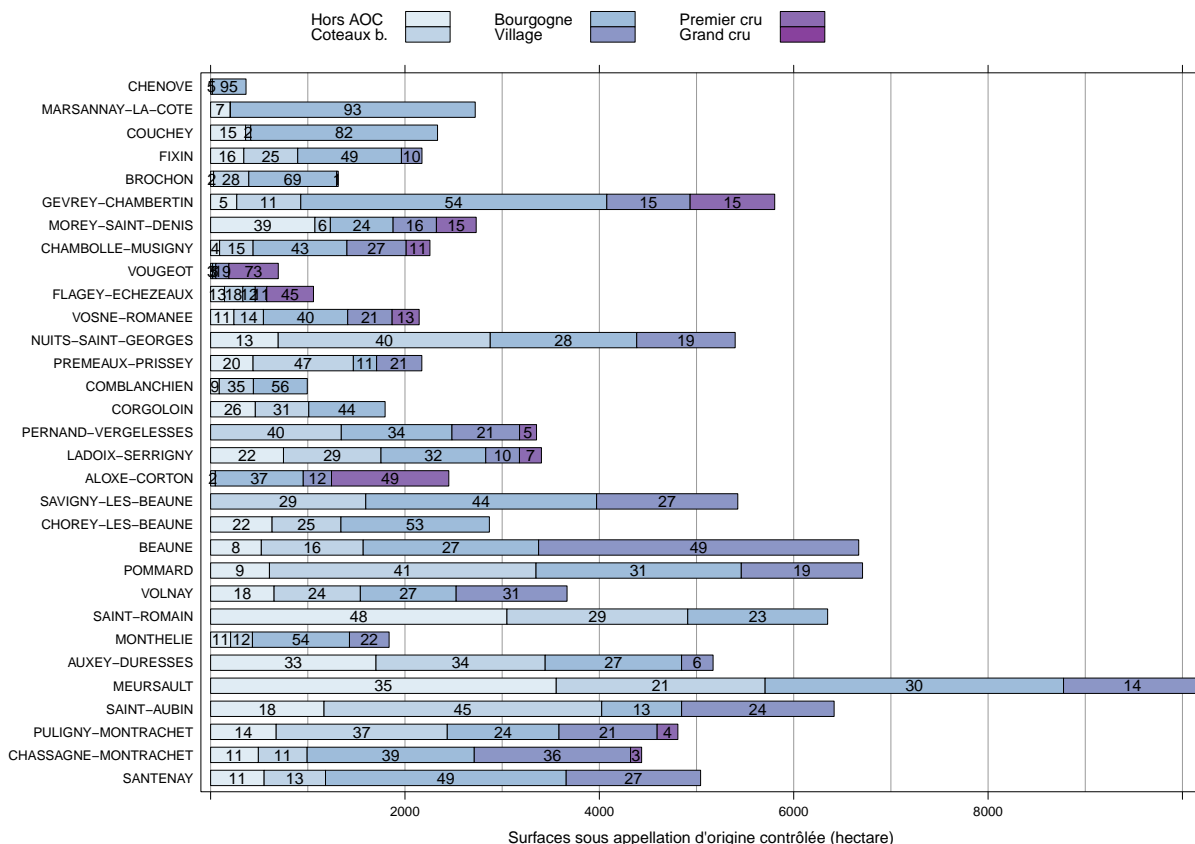
---

```
library(lattice) ; library(RColorBrewer)
fig.dat <- aggregate(model.matrix(~0+ factor(Reg.Ras$AOC))*
Reg.Ras$AREA/ 1000, by= list(Reg.Ras$LIBCOM), sum)
names(fig.dat) <- c("LIBCOM", "BGOR", "BOUR", "VILL", "PCRU", "GCRU")
fig.dat$LIBCOM <- factor(fig.dat$LIBCOM, lev= rev(levels(fig.dat$LIBCOM)))
fig.crd <- t(apply(fig.dat[, -1], 1, function(t) cumsum(t)- t/2))
fig.lab <- round(t(apply(fig.dat[, -1], 1, function(t) t/ sum(t)))* 100)
```

```

my.pal <- brewer.pal(n= 9, name = "BuPu")[ 2: 8]
barchart(LIBCOM~ BGOR+ BOUR+ VILL+ PCRU+ GCRU, xlim= c(-100, 10200),
  xlab="Surfaces sous appellation d'origine contrôlée (hectare)",
  data= fig.dat, horiz= T, stack= T, col= my.pal, border= "black",
  par.settings= list(superpose.polygon= list(col= my.pal)),
  auto.key= list(space= "top", points= F, rectangles= T, columns= 5,
    text=c("Hors AOC", "Coteaux b.", "Bourgogne",
      "Village", "Premier cru", "Grand cru")),
  panel=function(x, y, ...) {
    panel.grid(h= 0, v = -11, col= "grey60")
    panel.barchart(x, y, ...)
    ltext(fig.crd, y, lab= ifelse(fig.lab> 0, fig.lab, ""))})

```



### 3.3 Les AOC historiques

Hiérarchisation des données historiques par les nom de crus et s'il sont présents dans les nouvelles données.

### 3.4 Autre graphique

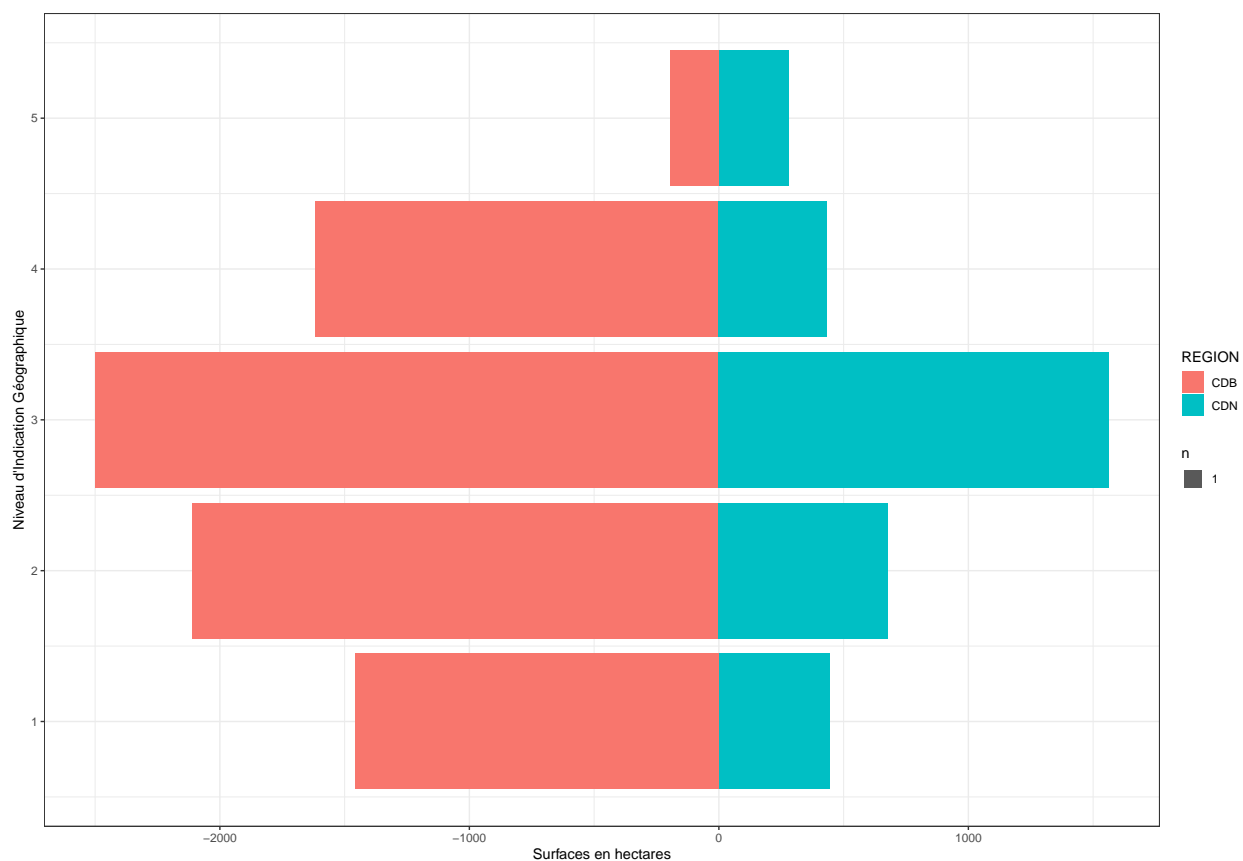
Définition de nos niveaux et implications en termes de surfaces sur la pyramides des AOC. Sur la Côte d'Or, on n'a pas vraiment la pyramide du BIVB. L'OP intègre mieux l'information, il ne faut pas mettre les 2 en concurrence. cette pratique est liée au principe de hiérarchisation des appellations d'origine, qui [...] s'emboîtent de manière pyramidale à partir d'une appellation régionale socle [...]. Dans cette optique, le vin

élaboré selon le cahier des charges d'une appellation hiérarchiquement supérieure répondrait de facto aux exigences de l'appellation régionale, dont les conditions de production sont moins contraignantes.

---

```
ddd <- aggregate(Reg.Ras$AREA/ 10000,  
                 by= list(Reg.Ras$AOC, Reg.Ras$REGION), sum, na.rm= TRUE)  
names(ddd) <- c("AOC", "REGION", "SURFACES")  
ddd$SURFACES[ddd$REGION== "CDB"] <- -ddd$SURFACES[ddd$REGION== "CDB"]  
library(ggplot2)  
ggplot(ddd, aes(x= AOC, y= SURFACES, fill= REGION))+  
  geom_bar(data= subset(ddd, REGION== "CDB"), stat= "sum") +  
  geom_bar(data= subset(ddd, REGION== "CDN"), stat= "sum") +  
  coord_flip() + theme_bw() + ylab("Surfaces en hectares") +  
  xlab("Niveau d'Indication Géographique")
```

---



### 3.5 yop

## 4 Modèle économétrique

### 4.1 Estimation du modèle

Modèle OGAM comme dans le papier compagnon, il est long à estimer, il est chargeable à partir des modèles estimées sur le serveur de l'INRA. Un préalable est le regroupement de variables géologique et pédologiques, au seuil de 1000 un peu arbitraire mais équilibré. Arbitrage entre XX et XX.

---

```
Reg.Ras$NOTATION <- factor(Reg.Ras$NOTATION)
tmp <- table(Reg.Ras$NOTATION)< 1000
Reg.Ras$GEOL <- factor(
  ifelse(Reg.Ras$NOTATION %in% names(tmp[ tmp]), "0AREF",
    as.character(Reg.Ras$NOTATION)))
Reg.Ras$NOUC <- factor(Reg.Ras$NOUC)
tmp <- table(Reg.Ras$NOUC)< 1000
Reg.Ras$PEDO <- factor(
  ifelse(Reg.Ras$NOUC %in% names(tmp[tmp]), "0AREF",
    as.character(Reg.Ras$NOUC)))
library(mgcv) ; load("Inter/gamodM.Rda")
## gam900 <- gam(AOC~ 0+ LIBCOM+ EXPO+ GEOL+ PEDO
##           + s(DEM)+ s(SLOPE)+ s(RAYAT)+ s(X, Y, k= 900)
##           , data= Reg.Ras, family= ocat(R= 5))
anova(gamodM$gam700)
```

---

Family: Ordered Categorical(-1,4.94,12.88,19.11)

Link function: identity

Formula:

AOC ~ 0 + LIBCOM + EXPO + s(DEM) + s(SLOPE) + s(RAYAT) + s(X,  
Y, k = listk[i]) + GEOL + PEDO

Parametric Terms:

	df	Chi.sq	p-value
LIBCOM	31	1724	<2e-16
EXPO	7	155	<2e-16
GEOL	14	407	<2e-16
PEDO	13	409	<2e-16

Approximate significance of smooth terms:

	edf	Ref.df	Chi.sq	p-value
s(DEM)	8.79	8.98	1014	<2e-16
s(SLOPE)	8.17	8.82	156	<2e-16
s(RAYAT)	7.97	8.74	649	<2e-16
s(X,Y)	666.56	683.49	78418	<2e-16

Effets des variables et pourcentages de bonnes prédictions.

## 4.2 Effets des variables biophysiques

La fonction par défaut permet de représenter graphiquement les relations statistiques entre l'altitude, la pente, le rayonnement solaire et la localisation et le rang dans la hiérarchie des AOC.

---

```
plot(gamodM$gam700, page= 1, scale= 0)
```

---

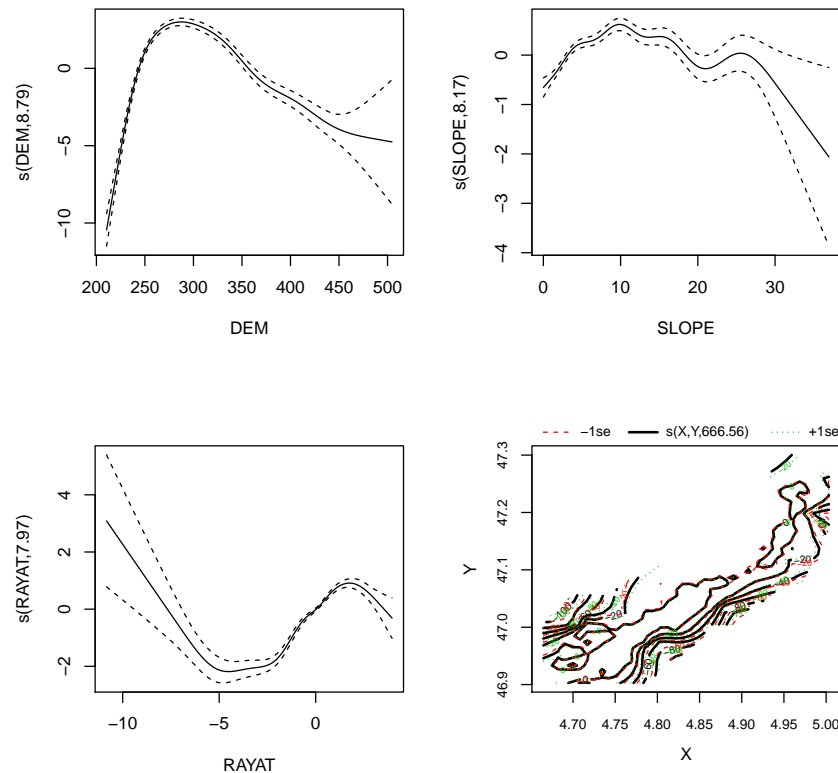


Figure 2: Effets semi-paramétriques de la topographie et de la localisation

## 4.3 Effets communaux

Christophe Lucand (dans [Wolikow and Jacquet \(2011\)](#)) cite les experts fondateurs (les mêmes qu'Olivier): Jullien, Morelot et Laval, supposent l'existence d'une hiérarchie commune en trois ou quatre catégories, avec au sommet les "têtes de cuvée" puis les premières cuvées. Puis il cite la thèse d'Olivier. A cette hiérarchie transversale se superpose une hiérarchie par villages qui ne détermine cependant en rien la réalité des zones d'approvisionnement concernées. Il s'agit plutôt d'identifications commerciales communes, investies d'un plus ou moins grand capital symbolique hérité. Ce capital symbolique hérité attribut un prestige plus ou moins grand à certaines communes ou propriétaires particulier.

---

```
library(latticeExtra) ; yop <- summary(gamodM$gam700)
```

---



```

plogi <- function(x) exp(x/ sqrt(2))/ (1+ exp(x/ sqrt(2)))
cf <- yop$p.coeff[ 4: 31]- mean(yop$p.coeff[ 4: 31])
se <- yop$se[ 4: 31]
zz <- data.frame(LIBCOM= substr(names(gamodM$gam700$coef[ 4: 31]), 7, 30),
  OS= 2* plogi(cf)- 1,
  OSi= 2* plogi(cf- 1.5* se)- 1,
  OSa= 2* plogi(cf+ 1.5* se)- 1)
segplot(reorder(factor(LIBCOM), OS)~ OSi+ OSa, length= 5, draw.bands= T,
  data= zz[order(zz$OS), ], center= OS, type= "o",
  unit = "mm", axis = axis.grid, col.symbol= "black", cex= 1,
  xlab= "Mesure de sup riorit  ordinaire et intervalles   10 %")

```

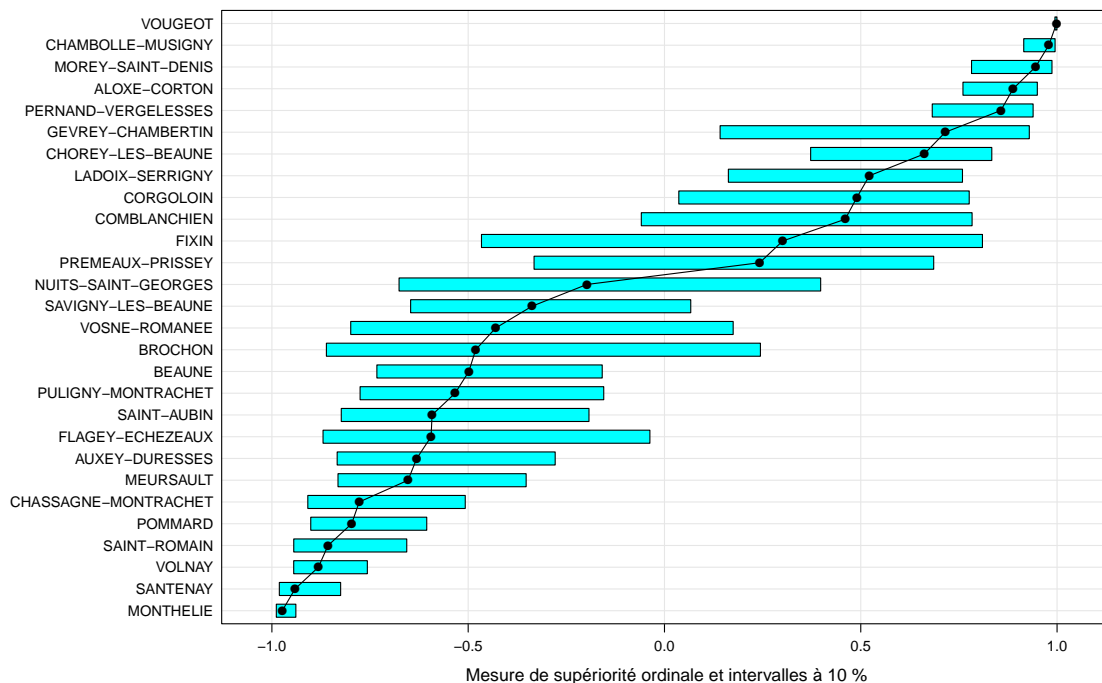


Figure 3: Effets communaux sur la classification AOC des parcelles

Mesure de sup riorit  ordinaire blabla. Attention, ce n'est pas la m me incertitude qui est repr sent e.

#### 4.4 Pr diction du score et classifications

```

load("Inter/gamod.Rda")
GAM900 <- predict(gamod$gam900, newdata= Geo.Ras@data, type= "terms")
Geo.Ras$G900raw <- rowSums(GAM900, na.rm= TRUE)
Geo.Ras$G900cor <- mean(GAM900[, 1], na.rm= T)+
  rowSums(GAM900[, -1], na.rm= TRUE)

load("Inter/modMontr.Rda")
tmp <- table(Reg.Ras$NOTATION)< 1000
Geo.Ras$GEOL <- factor(
  ifelse(Geo.Ras$NOTATION %in% names(tmp[ tmp]), "0AREF",

```

```

as.character(Geo.Ras$NOTATION)))

tmp <- table(Geo.Ras$NOUC)< 2200
Geo.Ras$PEDO <- factor(
  ifelse(!Geo.Ras$NOUC %in% names(tmp[tmp])),
    as.character(Geo.Ras$NOUC), "0AREF")
GAMONT <- predict(modMontr$gamm8, newdata= Geo.Ras@data, type= "terms")
Geo.Ras$GMONraw <- rowSums(GAMONT, na.rm= TRUE)
Geo.Ras$GMONcor <- mean(GAMONT[, 1], na.rm= T)+
  rowSums(GAMONT[, -1], na.rm= TRUE)

cor(Geo.Ras$G900cor, Geo.Ras$GMONcor)
cor(Geo.Ras$G900raw, Geo.Ras$GMONraw)

Geo.Aoc <- subset(Geo.Ras, PAOC== 1)
unini <- function(x) (x- min(x))/ (max(x)- min(x))
Geo.Aoc$G900raw <- round(unini(Geo.Aoc$G900raw)* 100, 2)
Geo.Aoc$G900cor <- round(unini(Geo.Aoc$G900cor)* 100, 2)
Geo.Aoc$GMONraw <- round(unini(Geo.Aoc$GMONraw)* 100, 2)
Geo.Aoc$GMONcor <- round(unini(Geo.Aoc$GMONcor)* 100, 2)

library(plotly)
## https://plot.ly/r/violin/

library(ggplot2)

## overlay
p <- ggplot(Geo.Aoc@data, aes(factor(AOC), G900raw))
p <- p + geom_violin(aes(colour= "#1268FF"), alpha= 0.3)
q <- p + geom_violin(aes(y= G900cor, colour= "#3268FF"), alpha= 0.3)
q+ ylim(40, 100)

p <- ggplot(Geo.Aoc@data, aes(factor(AOC), GMONraw))
p <- p + geom_violin(aes(colour= "#1268FF"), alpha= 0.3)
q <- p + geom_violin(aes(y= GMONcor, colour= "#3268FF"), alpha= 0.3)
q+ ylim(40, 100)

## split
library(plyr)
revAOC <- c("1"= "Coteaux b.", "2"= "Bourgogne", "3"= "Village",
  "4"= "Premier cru", "5"= "Grand Cru")
cc <- rbind(
  data.frame(AOC= revalue(factor(Geo.Aoc$AOC), revAOC),
    Score= Geo.Aoc$G900raw,
    Pr= "Uncorrected: without communes fixed effects"),
  data.frame(AOC= revalue(factor(Geo.Aoc$AOC), revAOC),
    Score= Geo.Aoc$G900cor,
    Pr= "Corrected: with communes fixed effects"))

ggplot(cc, aes(factor(AOC), Score, fill= Pr))+
  geom_split_violin()+
  ylab("100-Point Vineyard Quality Score")+
  ylim(40, 100)+ theme_minimal()+ xlab("")+
  geom_split_violin(draw_quantiles = c(0.25, 0.5, 0.75))+
  theme(legend.justification=c(0, 1), legend.position=c(0, 1),
    legend.title = element_blank())
#ggsave("Figures/MainPlot.png")

cc <- rbind(

```

```

data.frame(AOC= revalue(factor(Geo.Aoc$AOC), revAOC),
  Score= Geo.Aoc$GMONraw,
  Pr= "Uncorrected: without communes fixed effects"),
data.frame(AOC= revalue(factor(Geo.Aoc$AOC), revAOC),
  Score= Geo.Aoc$GMONcor,
  Pr= "Corrected: with communes fixed effects"))

ggplot(cc, aes(factor(AOC), Score, fill= Pr))+
  geom_split_violin()+
  ylab("100-Point Vineyard Quality Score")+
  ylim(40, 100)+ theme_minimal()+ xlab("")+
  geom_split_violin(draw_quantiles = c(0.25, 0.5, 0.75))+
  theme(legend.justification=c(0, 1), legend.position=c(0, 1),
    legend.title = element_blank())
#ggsave("Figures/MainPlot.png")

## PAS DE C-c C-c
GeomSplitViolin <- ggproto("GeomSplitViolin", GeomViolin,
  draw_group = function(self, data, ..., draw_quantiles = NULL) {
    # Original function by Jan Gleixner (@jan-glx)
    # Adjustments by Wouter van der Bijl (@Axeman)
    data <- transform(data, xminv = x - violinwidth * (x - xmin), xmaxv = x + violinwidth * (xmax - x))
    grp <- data[1, "group"]
    newdata <- plyr::arrange(transform(data, x = if (grp %% 2 == 1) xminv else xmaxv), if (grp %% 2 == 1) y else -y)
    newdata <- rbind(newdata[1, ], newdata, newdata[nrow(newdata), ], newdata[1, ])
    newdata[c(1, nrow(newdata) - 1, nrow(newdata)), "x"] <- round(newdata[1, "x"])
    if (length(draw_quantiles) > 0 & !scales::zero_range(range(data$y))) {
      stopifnot(all(draw_quantiles >= 0), all(draw_quantiles <= 1))
      quantiles <- create_quantile_segment_frame(data, draw_quantiles, split = TRUE, grp = grp)
      aesthetics <- data[rep(1, nrow(quantiles)), setdiff(names(data), c("x", "y")), drop = FALSE]
      aesthetics$alpha <- rep(1, nrow(quantiles))
      both <- cbind(quantiles, aesthetics)
      quantile_grob <- GeomPath$draw_panel(both, ...)
      ggplot2:::ggname("geom_split_violin", grid::grobTree(GeomPolygon$draw_panel(newdata, ...), quantile_grob))
    }
    else {
      ggplot2:::ggname("geom_split_violin", GeomPolygon$draw_panel(newdata, ...))
    }
  }
)

create_quantile_segment_frame <- function(data, draw_quantiles, split = FALSE, grp = NULL) {
  dens <- cumsum(data$density) / sum(data$density)
  ecdf <- stats::approxfun(dens, data$y)
  ys <- ecdf(draw_quantiles)
  violin.xminvs <- (stats::approxfun(data$y, data$xminv))(ys)
  violin.xmaxvs <- (stats::approxfun(data$y, data$xmaxv))(ys)
  violin.xs <- (stats::approxfun(data$y, data$x))(ys)
  if (grp %% 2 == 0) {
    data.frame(
      x = ggplot2:::interleave(violin.xs, violin.xmaxvs),
      y = rep(ys, each = 2), group = rep(ys, each = 2)
    )
  }
  else {
    data.frame(
      x = ggplot2:::interleave(violin.xminvs, violin.xs),
      y = rep(ys, each = 2), group = rep(ys, each = 2)
    )
  }
}

```

```

    }
  }

  geom_split_violin <- function(mapping = NULL, data = NULL, stat = "ydensity", position = "identity", ...,
                                draw_quantiles = NULL, trim = TRUE, scale = "area", na.rm = FALSE,
                                show.legend = NA, inherit.aes = TRUE) {
    layer(data = data, mapping = mapping, stat = stat, geom = GeomSplitViolin, position = position,
          show.legend = show.legend, inherit.aes = inherit.aes,
          params = list(trim = trim, scale = scale, draw_quantiles = draw_quantiles, na.rm = na.rm, ...))
  }

```

---

## 4.5 Aggrégation

```

gg <- aggregate(Geo.Aoc$G900raw, by= list(Geo.Aoc$AOC1b), mean, na.rm= T)
hh <- aggregate(Geo.Aoc$G900cor, by= list(Geo.Aoc$AOC1b), mean, na.rm= T)
head(gg[order(gg$x, decreasing= TRUE), ], n= 10)
head(hh[order(hh$x, decreasing= TRUE), ], n= 10)

gg <- aggregate(Geo.Aoc$GMONraw, by= list(Geo.Aoc$AOC1b), mean, na.rm= T)
hh <- aggregate(Geo.Aoc$GMONcor, by= list(Geo.Aoc$AOC1b), mean, na.rm= T)
head(gg[order(gg$x, decreasing= TRUE), ], n= 10)
head(hh[order(hh$x, decreasing= TRUE), ], n= 10)

gg[gg$Group.1== 'Romanée-Conti',]

```

---

## 5 Application cartographique

### 5.1 R code

#### AGGREGATION PAR LIEUX DITS

On utilise mapview, <https://r-spatial.github.io/mapview/>

- `sudo apt install libgdal-dev`
- `sudo ln -s /usr/lib/rstudio/bin/pandoc/pandoc /usr/local/bin`
- `webshot::installphantomjs()`

On pourrait également utiliser:

- <http://symbolixau.github.io/googleway/articles/googleway-vignette.html>
- <https://www.osgeo.org/projects/mapguide-open-source/>
- <http://geoserver.org/>
- <https://rstudio.github.io/leaflet/shiny.html>
- <https://github.com/mtennekes/tmap>

On peut mettre des graphiques quand on clique sur un polygone: [https://r-spatial.github.io/mapview/articles/articles/mapview\\_04-popups.html](https://r-spatial.github.io/mapview/articles/articles/mapview_04-popups.html)

also show info on the epsg code and the proj4string press and hold Ctrl and move the mouse. addMouseCoordinates also allows us to copy the info about the current mouse position to the clipboard by holding the Ctrl and left-clicking on the map.

---

```
pcgp <- function(x){
  sum(diag(table(cut(x$line, c(-Inf, x$family$getTheta(TRUE), Inf)),
    x$model[, 1])))/ nrow(x$model)* 100
}
pcgp(modMontr$gamm8)
summary(modMontr$gamm8)

plot(density((gam2$linear.pred- min(gam2$linear.pred))/
  (max(gam2$linear.pred)- min(gam2$linear.pred))))
prdat <- RegRank
prdat$LIBCOM <- "BROCHON"

gg <- predict(gam2, type= "response", newdata= prdat)
hh <- ifelse(gg[, 1]> 1- 1/1e16, 1- 1/1e16, gg[, 1])
prdat$score <- qlogis(1- hh)
RegRank$SCORE <- (prdat$score- min(prdat$score))/
  (max(prdat$score)- min(prdat$score))

plot(density(RegRank$SCORE))
library(plyr)
ee <- ddply(RegRank, .(CODEld),
  function(x) data.frame(Mean= mean(x$SCORE),
    Median= median(x$SCORE),
    WMean= weighted.mean(x$SCORE, x$Area)))
head(ee[order(ee$Mean, decreasing= TRUE), ], 20)

ff <- ddply(RegRank, .(LIBCOM),
  function(x) data.frame(Mean= mean(x$SCORE),
    Median= median(x$SCORE),
    WMean= weighted.mean(x$SCORE, x$Area)))
ff[order(ff$Mean, decreasing= TRUE), ]
ff[order(ff$WMean, decreasing= TRUE), ]

library(rgdal) ; library(rgeos) ; library(maptools)

NVA <- c("Coteaux b.", "Bourgogne", "Village", "Premier cru", "Grand cru")
names(NVA) <- 1: 5
library(plyr)
Geo.Aoc$NIVEAU <- revalue(factor(Geo.Aoc$AOC), NVA)
Geo.Aoc$NAME <- ifelse(Geo.Aoc$AOC== 5, as.character(Geo.Aoc$AOC1b),
  ifelse(Geo.Aoc$AOC< 4, as.character(Geo.Aoc$LIEUDIT), NA))
for (i in 1: nrow(Geo.Aoc)){
  if (is.na(Geo.Aoc$NAME[ i])){
    Geo.Aoc$NAME[ i] <- substr(Geo.Aoc$AOC1b[ i],
      regexpr(" cru+", Geo.Aoc$AOC1b[ i],
        perl= T)+ 5,
```

```

    } else {(Geo.Aoc$NAME[ i])}
  }
  nchar(as.character(Geo.Aoc$AOC1b[ i])))
}
Geo.Aoc$Concat <- paste0(Geo.Aoc$AOC, Geo.Aoc$LIBCOM, Geo.Aoc$NAME)
tmp_geo <- gBuffer(Geo.Aoc, byid= TRUE, width= 0)
Poly.tst <- unionSpatialPolygons(tmp_geo, Geo.Aoc$Concat)

Poly.tst$xx <- as.character(row.names(Poly.tst))
attrTb <- data.frame(xx= Geo.Aoc$Concat[ !duplicated(Geo.Aoc$Concat)],
  NIVEAU= Geo.Aoc$NIVEAU[ !duplicated(Geo.Aoc$Concat)],
  LIBCOM= Geo.Aoc$LIBCOM[ !duplicated(Geo.Aoc$Concat)],
  NAME= Geo.Aoc$NAME[ !duplicated(Geo.Aoc$Concat)])
Poly.tst <- merge(Poly.tst, attrTb, by= "xx")

Poly.tst$PrdRaw <- aggregate(Geo.Aoc$G900raw,
  by= list(Geo.Aoc$Concat), mean)[, 2]
Poly.tst$PrdCor <- aggregate(Geo.Aoc$G900cor,
  by= list(Geo.Aoc$Concat), mean)[, 2]
Poly.tst$Area <- aggregate(Geo.Aoc$AREA,
  by= list(Geo.Aoc$Concat), sum)[, 2]

Poly.tst@data[, 5: 7] <- round(Poly.tst@data[, 5: 7], 2)

library(RColorBrewer)
AocPal <- brewer.pal(5, "BuPu")
library(mapview)

mapviewOptions(basemaps = mapviewGetOption("basemaps")[ c(4, 1: 3, 5)])
tst <- mapview(Poly.tst,
  zcol= "NIVEAU", alpha.regions= .5,
  col.regions = AocPal,
  color= "white", legend.opacity= .5,
  label= Poly.tst$NAME, layerId= Poly.tst$xx,
  popup = popupTable(Poly.tst@data, feature.id= FALSE,
    zcol = c("NIVEAU", "LIBCOM", "NAME",
      "Area", "PrdRaw", "PrdCor")))

## addLogo(tst, "http://www7.inra.fr/fournisseurs/images/logo.jpg",
## width = 200, height = 100)
## mapviewOptions(basemaps = mapviewGetOption("basemaps")[ c(4, 1: 3, 5)],
## raster.palette = colorRampPalette(brewer.pal(9, "Greys")),
## vector.palette = colorRampPalette(brewer.pal(9, "YlGnBu")),
## na.color = "magenta",
## layers.control.pos = "topright")
tst
## create standalone .html
mapshot(tst, url = paste0(getwd(), "/DynMap/LieuxDits.html"))

## create .html and .png
mapshot(m, url = paste0(getwd(), "/DynMap/test.html"),
  file = paste0(getwd(), "/DynMap/test.png"),
  remove_controls = c("homeButton", "layersControl"))

```

---

## 5.2 Global

---

```
library(dygraphs)
```

```

library(shinydashboard)
library(shiny)
library(shinyjs)
library(mapview)
library(datasets)
library(dplyr)
library(leaflet)
#
latent <- Geo.Ras@data
#
# print(getwd())
# print(test)
#

my_map <-
  leaflet(Poly.Lxd) %>% addTiles() %>% addMouseCoordinates()

# #
# source("./ui.R")
# source("./server.R")
# #

shinyApp(ui, server)

```

---

## 5.3 Server

---

```

server <- shinyServer(function(input, output, session) {
  getData <- reactive({ latent[, as.character(input$prdd)] })
  output$myplot <- renderPlot({
    input$prdd
  }) # créer l'objet à plot => pour PlotOutput côté server
  output$myplot2 <- renderPlot({
    plot(density(getData()))
  }) # créer l'objet à plot => pour PlotOutput côté server
  output$mymap <- renderLeaflet({
    tst@map
  })
})

```

---

## 5.4 ui

---

```

ui <- dashboardPage(
  dashboardHeader(
    titleWidth = 150,
    title = "TITRE"),
  dashboardSidebar(width = 75,
    sidebarMenu(id = "tabs",
      menuItem("T1", tabName = "T1",
        icon = icon("angle-right ")),
      menuItem("T2", tabName = "T2",
        icon = icon("angle-right "))),
  dashboardBody(

```



```

tabItems(tabItem(tabName = "T1",
  fluidRow(
    box(width= 2,
      column(width = 12,
        selectInput("prdd",
          label = "Correction",
          choices=
            c("G900raw", "G900cor"),
          selected = 1))),
    box(width= 5,
      column(width = 12,
        plotOutput("myplot")))),
    # leafletOutput("mymap")))),
  tabItem(tabName = "T2",
    fluidRow(
      box(width = 5,
        column(width= 12,
          plotOutput("myplot2")))),
      fluidRow(
        box(width = 5,
          column(width = 12,
            plotOutput("myplot2"))))))))

```

---

## 6 Conclusion

Le chiffres d'affaire des signes de qualité c'est 32 milliards d'euros et le budget de l'INAO 32 millions d'euros, c'est un millièmè du chiffre d'affaires.

---

```
sessionInfo()
```

---

## 7 Bibliographie

### References

- Ay, J.-S. (2011). Hétérogénéité de la terre et rareté économique. Theses, Université de Bourgogne.
- Capus, J. (1947). *L'Évolution de la législation sur les appellations d'origine. Genèse des appellations contrôlées*. L. Larmat.
- Coestier, B. and Marette, S. (2004). *Economie de la qualité*. La découverte.
- Garcia, J.-P. (2011). *Les climats du vignoble de Bourgogne comme patrimoine mondial de l'humanité*. Ed. Universitaires de Dijon.
- Humbert, F. (2011). L'INAO, de ses origines à la fin des années 1960: genèse et évolutions du système des vins d'AOC. Ph.D. thesis, Université de Bourgogne.
- Wolikow, S. and Jacquet, O. (2011). *Territoires et terroirs du vin du XVIIIe au XXIe siècles*. Éditions Universitaires de Dijon.

## A Annexes

---

```
as.vector(Geo.CDem$IDU[Geo.CDem$AOC== "BGOR" & rowSums(Geo.CDem@data[, c(24, 26: 29)])> 0])
as.vector(Geo.CDem$IDU[Geo.CDem$AOC== "BOUR" & rowSums(Geo.CDem@data[, 26: 29])> 0])
as.vector(Geo.CDem$IDU[Geo.CDem$AOC== "VILL" & rowSums(Geo.CDem@data[, c(26, 29)])> 0])
as.vector(Geo.CDem$IDU[Geo.CDem$AOC== "PCRU" & Geo.CDem@data[, 26]> 0])
```

---

```
[1] "21412000AZ0139" "21464000AN0094" "21492000AR0011"
[4] "21492000BN0045" "215690000C0840"
```

```
[1] "210370000A0507" "211100000AK0116" "211500000AM0096"
[4] "21428000AA0019" "21582000BC0069"
```

```
[1] "210370000AH0094" "210370000AH0096" "211100000AM0101"
[4] "211330000AB0401" "211330000AC0005" "211330000AC0003"
[7] "211330000AC0002" "211330000AC0004" "215120000AE0292"
[10] "215820000AL0049"
```

```
[1] "214420000AB0315"
```

Pour retrouver les intitulés des UCS, nous utilisons le site web <https://bourgogne.websol.fr/carto> où les différents types de sols qui composent les UCS sont consultables. Le travail manuel a consisté à extraire les coordonnées Lambert 93 d’au moins une parcelle par UCS et d’aller chercher sur le site le nom de l’UCS correspondante. Nous voyons également que lorsque l’UCS est un numéro manquant c’est qu’il s’agit de sols artificialisés (Chenôve, Nuits et Beaune). Il y a un léger effet frontière au sud sur les valeurs qui ne sont pas appariées.

---

```
yy <- data.frame(coordinates(GCDtmp3), GCDtmp3$NOUC)
yy[!duplicated(GCDtmp3$NOUC), ]
plot(GCDtmp3)
plot(GCDtmp3[GCDtmp3$NOUC== "0",], col= "blue", add= T, pch= 20)
```

---