

Calcul d'allocations de l'enquête pratiques culturelles en viticulture 2019

Maxime SIMONOVICI

Insee, Direction Régionale Centre-Val de Loire

JMS 2022 - Session 24 - 31 mars 2022



École nationale
de la statistique
et de l'analyse
de l'information



- 1 Plans de sondage précédents et nouveaux objectifs
- 2 Calculs empiriques d'allocations
- 3 Résultats, conclusions et perspectives

1-1 Présentation des enquêtes pratiques culturelles en viticulture

Objectifs des enquêtes pratiques culturelles :

- suivre le plan Ecophyto, répondre à une directive européenne ;
- collecter des données sur les travaux des agriculteurs.

L'indicateur de fréquence de traitement (IFT) d'une parcelle culturale c :

$$IFT(c) = \sum_{t \in c} IFT_t = \sum_{t \in c} \frac{DA_t}{DR_t} \cdot PST_t$$

Un indicateur :

- complexe, divisible, territorialisé ;
- qui constitue un outil d'évaluation de politiques publiques ;
- avec une précision visée sur sa moyenne de 2,5 %.

1-2 Historique des plans de sondage précédents

Informations générales :

- **base de sondage** : CVI (Casier viticole informatisé) ou CIVC (Comité interprofessionnel du vin de Champagne) pour le bassin champenois ;
- **unité de tirage** : la sous-parcelle cadastrale p ;
- **unité enquêtée** : la parcelle culturale c dont relève la sous-parcelle cadastrale p tirée : un ensemble contigu de vignes d'un seul cépage et conduit selon des pratiques homogènes (fertilisation, traitements phytosanitaires, etc.) ;
- **unité de collecte** : l'exploitation agricole (une même exploitation peut être enquêtée pour plusieurs parcelles).

Compte tenu du nombre important de sous-parcelles cadastrales, il s'avère peu probable qu'une parcelle culturale soit tirée plus d'une fois.

1-2 Historique des plans de sondage précédents

Dès la première enquête PKViti2006 :

- une unité statistique particulière : la parcelle culturale c non identifiable dans la base de sondage ;
- une stratification selon le département et type de vin ;
- un tirage à taille fixe avec des probabilités de tirage inégales proportionnelles à leur taille.

La probabilité d'inclusion π_c d'une parcelle culturale c appartenant à la strate S_h vaut alors :

$$\pi_c \approx \sum_{p \in c} \pi_p = \sum_{p \in c} n_h \cdot \frac{\text{surf}(p)}{\sum_{p \in S_h} \text{surf}(p)} = n_h \cdot \frac{\text{surf}(c)}{\sum_{p \in S_h} \text{surf}(p)}$$

avec p la ou les sous-parcelle(s) cadastrale(s) associée(s) à la parcelle culturale c .

1-2 Historique des plans de sondage précédents

Depuis 2010, un panel interrogé à trois reprises avec :

- un maintien du lien entre la sous-parcelle cadastrale et la parcelle culturale ;
- un calcul similaire des probabilités d'inclusion ;
- une somme des coefficients d'extrapolation égale à la surface du bassin ;
- une non-réponse faible (5 % maximum).

Quelques transformations et limites :

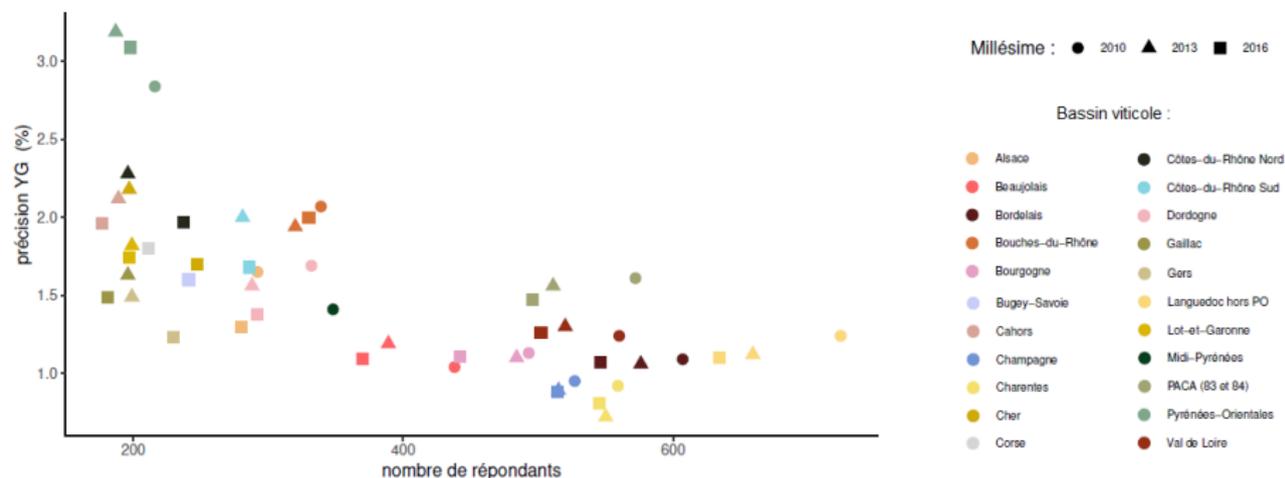
- l'intégration successive de nouveaux bassins viticoles ;
- une stratification par type de vin pas toujours appropriée ;
- des calages différents selon les enquêtes (60 zones de calage en 2016) ;
- une dispersion des coefficients d'extrapolation plus importante en 2016.

1-3 Historique des méthodes de calcul des allocations

- en 2006 : calcul des allocations n_h en recherchant un taux de sondage uniforme dans chaque strate S_h selon leur superficie (en ha) ;
- à partir de 2010 pour les nouveaux bassins :
 - en fonction d'un abaque indiquant la taille de l'échantillon nécessaire pour obtenir une précision donnée (régression linéaire simple tous bassins confondus) ;
 - des différences du nombre minimum de questionnaires (350, 200 puis 250).

1-3 Historique des méthodes de calcul des allocations

Un lien entre nombre de répondants et précision de l'IFT moyen différent entre les bassins viticoles :



Sources : Agreste - Enquêtes sur les pratiques phytosanitaires (2010, 2016) et culturales (2013) en viticulture.

1-4 Objectifs et contexte de la préparation du plan de sondage de PKViti2019

Objectifs :

- maintenir une précision de 2,5 % sur la moyenne de l'IFT ;
- prendre en compte une nouvelle demande : estimer l'IFT moyen des parcelles menées en Agriculture Biologique (AB) ;
- maîtriser les coûts.

Des premiers essais :

- plan de sondage à deux degrés (exploitation puis parcelle) ;
- calculs théoriques d'allocations.

Solution retenue :

adapter partiellement le plan de sondage actuel et affiner le calcul des allocations

- 1 Plans de sondage précédents et nouveaux objectifs
- 2 Calculs empiriques d'allocations**
- 3 Résultats, conclusions et perspectives

2-1 Méthode de Bootstrap : rappels théoriques

L'algorithme général :

- 1 Pour b allant de 1 à B (le nombre d'itérations choisi) :
 - 1 tirer un échantillon i.i.d. de taille n selon la loi \hat{F} conditionnellement à l'échantillon de départ X_1, \dots, X_n ;
 - 2 calculer $\hat{\theta}_b$;

- 2 Calculer l'estimateur Bootstrap de θ : $\widehat{\theta}_{Boot} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b$;

- 3 Calculer l'estimateur Bootstrap de la variance de θ :

$$\widehat{V}_{Boot} = \frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b - \widehat{\theta}_{Boot})^2$$

En mesurer la qualité via plusieurs indicateurs (biais relatif, erreur quadratique moyenne, coefficient de variation des estimations Bootstrap de variance).

2-2 Objectifs des méthodes de Bootstrap utilisées

Objectif principal de ces travaux empiriques :

dans chaque bassin viticole : prévoir l'impact d'une baisse de l'allocation sur la précision de l'IFT moyen

Plusieurs essais successifs de Bootstrap sur PhytoViti2016 :

- directs (sous ou sur-échantillonnage) puis de type pseudo-population ;
- à façon pour intégrer des propriétés du plan de sondage de l'enquête.

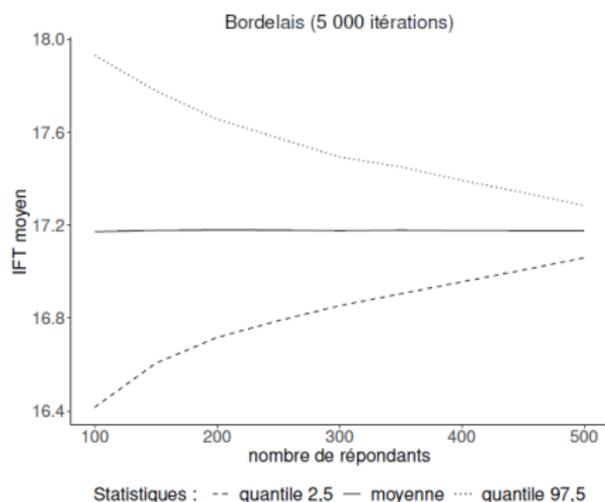
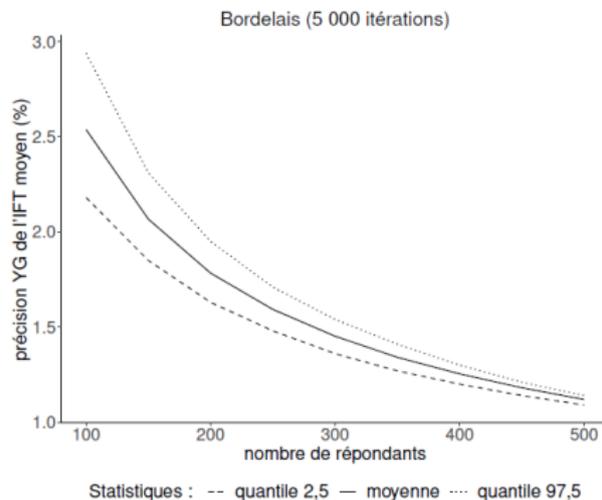
Spécificités de l'algorithme de la méthode retenue (Bootstrap à façon) :

- calculer proportionnellement dans chaque croisement $g = \text{Strate} \otimes \text{zone de calage}$, t_g le nombre de parcelles à tirer ;
- recalculer le poids $COEF2(c)$ en $w_2(c)$ de tous les individus de l'échantillon afin de garantir l'égalité :

$$\sum_{c \in s_b} w_2(c) \cdot \mathbb{1}_{c \in g \cap s_b} = \text{surf}(g)$$

2-3 Programmation et résultats

- calculer à chaque itération b via une fonction en R (CalvaR) : $\hat{\theta}_b$ l'estimateur Bootstrap propre à b ainsi que les variances et précisions au sens d'Horvitz-Thompson et de Yates-Grundy ;
- produire le graphique et le tableau des données associées des précisions :



Source : Agreste - Enquête sur les pratiques phytosanitaires en viticulture 2016.

2-4 Affinement des calculs d'allocation

Sous certaines hypothèses, des enseignements et précautions :

- dans les bassins non stratifiés : la baisse des allocations est possible tout en maintenant une précision satisfaisante ;
- ne pas diminuer les allocations autant que les résultats ne le suggèrent.

En parallèle, le plan de sondage est réorienté :

- la stratification par type de vin est abandonnée ;
- la détermination de l'IFT moyen des parcelles menées en Agriculture Biologique (AB) est prévue dans un nombre limité de bassins ;
- une liste de bassins candidats est ainsi constituée selon certains critères : surface en AB significative, nombre suffisant d'exploitations et donc de parcelles menées en AB (pour limiter les multi-enquêtées), un appariement des fichiers CVI – Agence Bio de qualité satisfaisante.

2-4 Affinement des calculs d'allocation

Une nouvelle stratification est mise en place selon le mode de conduite des parcelles (conventionnelle, biologique).

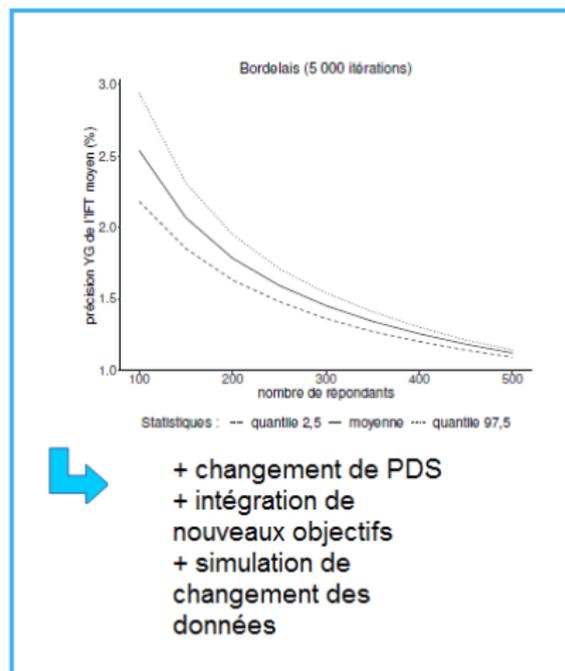
De nouvelles méthodes de Bootstrap à façon sont testées dans les bassins stratifiés pour simuler le plan de sondage de PKViti2019 :

- en faisant varier les allocations des différentes strates ;
- sous certaines contraintes : $n_{conv} \geq n_{AB}$, $q_{97,5}(prec_{AB}) \leq 3,5 \%$ et $q_{97,5}(prec_{bassin}) \leq 2,5 \%$;
- en faisant varier la dispersion des IFT des parcelles menées en AB tirées à chaque itération.

Des derniers ajustements sont effectués avant la détermination des allocations officielles.

2-5 Résumé de la méthode

Taille de l'échantillon	Précision attendue de l'IFT (en %)
100	3,09
150	2,9
200	2,72
250	2,53
300	2,34
350	2,16
400	1,97
450	1,79
500	1,6



Plan

- 1 Plans de sondage précédents et nouveaux objectifs
- 2 Calculs empiriques d'allocations
- 3 Résultats, conclusions et perspectives**

3-1 Collecte et précisions des données diffusées

Une collecte marquée par une part de non-réponse plus forte (88,2% de questionnaires exploitables).

Les objectifs de précision atteints dans la majorité des strates et bassins :

- $\leq 2,5$ % dans 18 bassins viticoles sur 22 ;
- avec une baisse considérable des allocations (notamment dans les bassins les plus importants) ;
- une diffusion de l'IFT moyen des parcelles menées en AB pour la première au niveau de certains bassins viticoles.

Quelques limites et approfondissements nécessaires :

- le cas particulier du bassin des Pyrénées-Orientales ;
- un seul bassin où l'objectif a été loin d'être atteint (Côtes-du-Rhône Sud, 3,8%).

3-2 Conclusions et perspectives

Conclusions générales :

- des calculs empiriques d'allocations encourageants malgré une collecte perturbée ;
- une méthode permettant d'utiliser de manière plus développée les données des enquêtes précédentes afin de viser de nouveaux objectifs.

Perspectives :

- des améliorations possibles du plan de sondage (ex : utilisation de nouvelles informations, changement de tirage) ;
- la richesse des nouvelles données en AB pour affiner les calculs d'allocations de la prochaine enquête.

Merci pour votre attention.