IVES Technical Reviews
vigne & vin

https://doi.org/10.20870/IVES-TR.2024.8212

Comment mieux estimer le nombre de grappes à la parcelle ?

Baptiste Oger[1*], Cécile Laurent[2], Philippe Vismara[3], Bruno Tisseyre[1]

[1]ITAP, Univ. Montpellier, L'institut Agro Montpellier, INRAE, France
[2]Fruition Sciences, MIBI, 672 Rue du Mas de Verchant, 34000 Montpellier, France
[3]MISTEA, Univ. Montpellier, L'institut Agro Montpellier, INRAE, France

En viticulture, l'estimation en début de saison du nombre moyen de grappes de raisin par pied est cruciale pour la planification, les investissements et le marketing d'une entreprise viticole[1]. Bien que l'échantillonnage soit largement pratiqué pour estimer le nombre de grappes de raisin par pied, il n'existe pas de protocole d'échantillonnage clairement établi qui puisse servir de référence lors de la réalisation de ces estimations. Chaque praticien a donc son propre protocole d'échantillonnage. Dans ce travail, l'effet des différences entre les protocoles d'échantillonnage en termes d'erreurs d'estimation a été étudié afin de donner quelques recommandations sur les meilleures pratiques d'échantillonnage possibles.

Les différences entre les protocoles d'échantillonnage

Des praticiens de plusieurs organisations ont été interrogés sur les pratiques d'échantillonnage qu'ils appliquent pour estimer le nombre moyen de grappes par vigne, et des variations considérables ont été observées. Les principales différences concernent :

- Le protocole de comptage, qui peut inclure ou omettre les vignes manquantes.
- La disposition des vignes échantillonnées, qui peuvent être regroupées dans des zones d'échantillonnage de taille variable (un nombre variable de vignes consécutives échantillonnées ensemble le long d'un rang).
- Le nombre total de vignes échantillonnées par parcelle.

Les sections qui suivent fournissent aux praticiens quelques recommandations basées sur les résultats expérimentaux concernant l'estimation du nombre moyen de grappes par vigne à la parcelle.

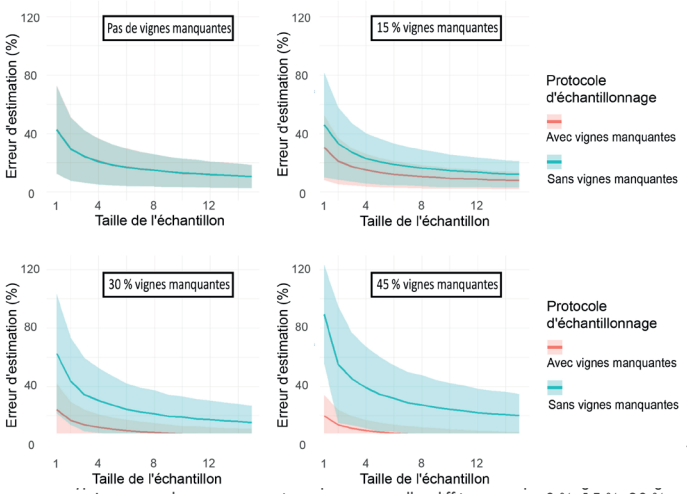
L'estimation doit être axée sur le nombre de grappes et ignorer les vignes manquantes

Lors de l'estimation par échantillonnage, la taille de l'échantillon doit être adaptée à la variable d'intérêt. Par conséquent, deux variables ayant des propriétés différentes doivent être estimées à l'aide de protocoles d'échantillonnage distincts. En viticulture, par exemple, si l'on suit un protocole d'échantillonnage dans lequel les vignes manquantes sont comptées comme des vignes à 0 grappe, deux composantes du rendement sont estimées simultanément : la proportion de vignes manquantes et le nombre de grappes. Toutefois, étant donné la nature intrinsèquement distincte de ces deux composantes du rendement, elles doivent être échantillonnées indépendamment selon des protocoles spécifiques.

Pour les quatre parcelles de la Figure 1, l'erreur d'estimation diminue logiquement à mesure que la taille de l'échantillon (nombre de vignes effectivement échantillonnées) augmente, quel que soit le protocole d'échantillonnage. La figure montre également qu'avec l'augmentation du nombre de vignes manquantes et mortes sur une parcelle, leur prise en compte lors de l'échantillonnage du nombre moyen de grappes par pied (en bleu) conduit à des erreurs d'estimation plus élevées. Ce phénomène illustre la nécessité d'estimer indépendamment chacune des deux composantes du rendement pour réduire les erreurs d'estimation, en particulier lorsque la proportion de vignes manquantes au sein des parcelles est importante. Par souci de clarté, les parcelles présentées ici sont théoriques, mais les résultats sont cohérents avec ceux obtenus dans des parcelles réelles[2 3].

Les observations doivent être réparties sur plusieurs zones de mesure

Les composantes du rendement sont souvent organisées dans l'espace, ce qui signifie que des vignes proches les unes des autres sont plus susceptibles de présenter des propriétés similaires que des vignes plus éloignées. Dans le cas du nombre de grappes par pied de vigne, deux pieds proches ont plus de chances d'avoir un nombre de grappes similaire. En échantillonnant des vignes consécutives sur une seule zone de mesure, on risque de surestimer (ou de sous-estimer) le nombre de grappes à la parcelle si cette zone présente un peu plus



manquantes. Les protocoles sont comparés sur quatre parcelles différentes avec 0 %, 15 %, 30 % ou 45 % de vignes manquantes (données simulées).

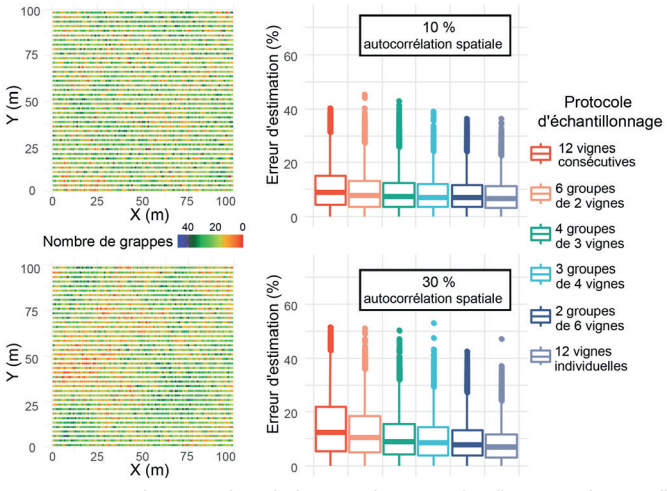


FIGURE 2. Erreur d'estimation liée à la disposition des vignes échantillonnées sur deux parcelles (données simulées).

Ce travail est disponible sous la licence Creative Commons Attribution 4.0 International.

(ou un peu moins) de grappes que la moyenne. Il est donc préférable de répartir les observations sur plusieurs zones de mesure, au moins deux ou trois, pour limiter ce risque.

La Figure 2 représente deux parcelles. Sur la parcelle du haut, le nombre de grappes par pied est faiblement organisé dans l'espace : des vignes ayant beaucoup de grappes coexistent avec des vignes proches ayant peu de grappes. Sur la parcelle du bas, le nombre de grappes par pied est plus structuré : dans une zone proche, la plupart des vignes ont des nombres de grappes similaires. Sur chaque parcelle, six protocoles d'échantillonnage répartissant 12 vignes sur 1, 2, 3, 4, 6 ou 12 zones de mesure ont été comparés.

Sur les deux parcelles, les erreurs d'estimation sont plus élevées lorsque les vignes sont regroupées sur une seule zone de mesure (boîte à moustaches rouge) que lorsqu'elles sont regroupées sur au moins deux zones de mesure. La Figure 2 illustre comment, pour un nombre égal de vignes observées, plus les vignes échantillonnées sont réparties sur un grand nombre de zones de mesure, moins les erreurs d'estimation sont importantes. Ce phénomène est encore plus prononcé lorsque le nombre de grappes par pied est autocorrélé dans l'espace, comme le montrent les erreurs d'estimation plus importantes sur la parcelle présentant 30 % d'autocorrélation spatiale.

Adapter la taille de l'échantillon à la variabilité observée

La dernière partie de cette revue technique traite de la relation entre l'erreur d'estimation, la taille de l'échantillon et la variabilité de l'échantillon. La variabilité de l'échantillon est exprimée ici sous la forme du coefficient de variation (CV). Ce coefficient est simplement calculé en divisant l'écart-type de l'échantillon par sa moyenne (Eq. 1) :

$$CV = \frac{écart - type}{moyenne}$$

Il s'agit d'une méthode normalisée pour quantifier la diversité des valeurs au sein de l'échantillon. Les détails statistiques de la méthode figurent dans l'article associé à cette revue technique[2].

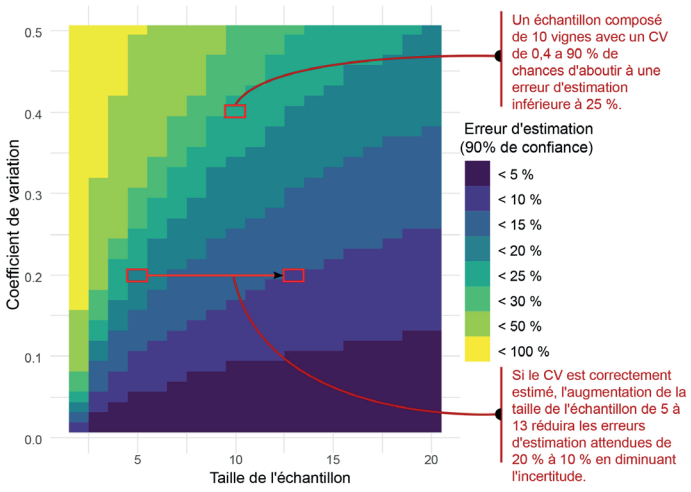


FIGURE 3. Erreur d'estimation attendue d'un échantillon en fonction de sa taille (nombre de vignes échantillonnées) et de son coefficient de variation.

La Figure 3 illustre comment, pour une taille d'échantillon donnée, les erreurs d'estimation varient en fonction du coefficient de variation (CV) de l'échantillon. Elle montre également comment la variabilité (représentée par le CV) des observations faites sur quelques vignes peut être utilisée pour définir une taille d'échantillon appropriée afin d'atteindre l'erreur d'estimation attendue. Moins la variabilité des observations au sein d'un échantillon est importante, plus la confiance en l'estimation est élevée. Un échantillon très variable (CV élevé) reflète une hétérogénéité importante à la parcelle, laquelle est associée à un risque d'erreur élevé. Par conséquent, lorsqu'une forte variabilité est observée au cours de l'échantillonnage, il est conseillé d'augmenter la taille de l'échantillon afin d'atténuer les erreurs d'estimation élevées.

En outre, l'augmentation de la taille de l'échantillon permet d'obtenir des gains plus importants lorsque la variabilité de l'échantillon est élevée. Sous réserve que le CV de l'échantillon ne change pas, l'augmentation du nombre de vignes échantillonnées de 5 à 13 (Figure 3) réduit la limite supérieure de l'intervalle de confiance pour les erreurs d'estimation de ~39 % à ~22 % si le CV est de 0,4, et seulement de ~19 % à ~11 % si le CV est de 0,2.

En calculant son CV et en utilisant les informations fournies par la Figure 3, il est possible d'évaluer en temps réel la qualité d'un échantillon. Le praticien peut alors décider de poursuivre l'échantillonnage en identifiant la taille de l'échantillon qui, à coefficient de variation égal, donnerait la précision souhaitée. Cette méthode pourrait également aider à détecter les parcelles où la taille de l›échantillon nécessaire pour atteindre un niveau d›erreur souhaité n›est pas réalisable, ce qui permettrait d›arrêter l›échantillonnage et de concentrer les efforts d›estimation sur d›autres parcelles. Il convient de souligner que si les nouvelles mesures modifient le coefficient de variation, il sera alors nécessaire de se déplacer verticalement sur le graphique pour tenir compte de ce changement. Pour faciliter la lecture, une forme tabulaire de cette figure est disponible gratuitement en annexe (voir référence Zenodo[4]). Cette annexe introduit également des niveaux de confiance de 99 %, 95 %, 75 % et 50 % pour mieux répondre à la confiance souhaitée.

Conclusions

La présente étude met en évidence ce qu'il convient de faire ou de ne pas faire lors de l'estimation du nombre moyen de grappes par pied de vigne par le biais d'un échantillonnage à la parcelle. La première recommandation concerne la mise en œuvre d'un protocole d'échantillonnage spécifique à chaque composante du rendement. Bien qu'il puisse être tentant d'échantillonner simultanément le nombre de pieds manquants et le nombre moyen de grappes par pied, ces deux composantes ont des propriétés différentes et requièrent donc chacune des protocoles spécifiques. La deuxième recommandation concerne l'organisation des zones de mesure au sein d'une parcelle. Bien qu'il soit possible de les regrouper en zones d'échantillonnage composés de plusieurs vignes consécutives pour tenir compte des contraintes opérationnelles, il devrait y avoir un minimum de 2 ou 3 zones de mesure pour éviter d'augmenter les erreurs d'estimation. Enfin, ce travail démontre que la définition d'une taille d'échantillon unique appliquée uniformément à toutes les parcelles est contre-productive. Au contraire, la taille de l'échantillon doit être déterminée en fonction de l'hétérogénéité observée, en choisissant le meilleur compromis entre l'erreur d'estimation et l'effort d'échantillonnage, en termes de temps et de coûts. Dans le cas où le coefficient de variation peut être calculé en temps réel pendant l'échantillonnage, les moyens décrits dans cet article pourraient être utilisés pour déterminer la précision associée à un échantillon et pour adapter les protocoles en cours de route. ■

Remerciements : *Ce travail a été financé par la région Occitanie.*

1 Laurent, C., Oger, B., Taylor, J. A., Scholasch, T., Metay, A., & Tisseyre, B. (2021). A review of the issues, methods and perspectives for yield estimation, prediction and forecasting in viticulture. *European Journal of Agronomy, 130*, 126339. https://doi.org/10.1016/j.eja.2021.126339

2 Oger, B., Laurent, C., Vismara, P., & Tisseyre, B. (2023a). How to better estimate bunch number at vineyard level? *OENO One, 57*(3), 27-39. https://doi.org/10.20870/oeno-one.2023.57.3.7404

3 Oger, B., Zhang, Y., Gras, J. P., Valloo, Y., Faure, P., Brunel, G., & Tisseyre, B. (2023b). High spatial resolution dataset of grapevine yield components at the within-field level. *Data in Brief, 50*, 109580. https://doi.org/10.1016/j.dib.2023.109580

4 Oger, B., Laurent, C., Vismara, P., & Tisseyre, B. (2024). Sampling to estimate the average number of bunches per vine: table of confidence intervals for estimation errors. *Zenodo*. https://doi.org/10.5281/zenodo.10594995