

UN EXEMPLE D'UTILISATION DES DONNÉES « AGRICULTEUR »

# ESTIMER LES RENDEMENTS des variétés de blé tendre

François Piraux - [f.piraux@arvalis.fr](mailto:f.piraux@arvalis.fr)

Afin d'améliorer l'évaluation variétale du blé tendre d'hiver, Arvalis a étudié la cohérence statistique de données de rendement collectées par les coopératives Terrena et Axérial et de données d'essais « Variétés » en micro-parcelles. Les résultats sont concluants.



Deux jeux de données de rendement de variétés de blé, l'un provenant d'essais et l'autre de données « agriculteur », ont été analysés. Leur classement respectif des variétés selon leur productivité est cohérent à 86 %.

**D**epuis de nombreuses années, la recherche en agriculture repose, au moins en partie, sur l'utilisation d'expérimentations en micro-parcelles.

Les principes de base de l'expérimentation (*encadré*), attribués à Ronald Fisher au début du XX<sup>e</sup> siècle, offrent un cadre rigoureux permettant la mise en évidence de relations de cause à effet. La connais-

sance ainsi produite peut ensuite être transposée à l'échelle d'une parcelle, afin d'aider les agriculteurs à prendre les meilleures décisions possibles pour conduire leurs parcelles.

## L'EXPÉRIMENTATION AGRONOMIQUE EN TROIS MOTS



Les principes de base des essais en micro-parcelles permettent de mettre en évidence des relations de cause à effet.

© Altélic

### RANDOMISATION

Elle consiste à injecter une part d'aléatoire dans la répartition des modalités expérimentales dans un essai. Elle a pour but d'éviter les biais.

### RÉPÉTITION

Les répétitions sont nécessaires pour estimer la variabilité expérimentale et pour augmenter la précision des estimations des moyennes des modalités étudiées.

### BLOCKING

Il consiste à répartir les modalités expérimentales au sein de blocs de micro-parcelles les plus homogènes possibles, dans le but de contrôler une partie de la variabilité expérimentale.

Avec l'essor du numérique, des données obtenues à d'autres fins que la recherche sont rendues accessibles. Les données « agriculteur » stockées dans les bases de données des logiciels de gestion parcellaire sont ainsi une source potentiellement importante d'information. Ces données peuvent-elles être valorisées pour répondre à des questions scientifiques ? Arvalis, en partenariat avec les coopératives Terrena et Axérial, a étudié leur intérêt pour enrichir l'évaluation variétale du blé tendre d'hiver.

26 799 données de rendement, issues d'agriculteurs répartis sur 1698 communes, ont été collectées auprès de Terrena et Axérial en 2015, 2017 et 2018. Elles ont été « confrontées » à 4291 données expérimentales collectées sur 74 communes différentes, provenant d'essais « Variétés » réalisés par les deux coopératives et Arvalis durant les mêmes années et sur la même zone géographique (figure 1 p.58).

Dans un premier temps, les deux jeux de données ont été analysés indépendamment afin d'estimer pour chacun les effets « variété » sur le rendement. L'écart entre le rendement moyen d'une variété et la moyenne générale des rendements a ainsi été



Céréales

Blé tendre d'hiver

# Greka

La qualité est sa meilleure arme



COTATION  
CTPS

102.8%

en % des témoins CTPS Sud  
de 2019/20\*

- BPS - VOp<sup>Extensible</sup>
- PS et protéines élevés, GPD+
- Excellent profil maladies ++
- Potentiel de rendement élevé

Retrouvez plus d'informations sur notre site internet :  
[www.agriobtentions.fr](http://www.agriobtentions.fr)



calculé pour chaque variété, avec chaque jeu de données.

Puis une analyse globale a été effectuée sur les données des deux réseaux, afin de voir si leurs effets « variétés » étaient corrélés. Les données agrégées ont été analysées au moyen d'un modèle mixte prenant en compte des effets « variété » et « type de données » (agriculteurs ou expérimentales), ainsi qu'une éventuelle interaction entre le type de données et les variétés.

### Des données « agriculteur » plus variables, mais plus nombreuses

L'analyse séparée des deux jeux de données a permis notamment de déterminer la précision des estimateurs des effets « variété » en fonction du nombre de données disponibles par variété. Les données expérimentales conduisent à des estimations plus précises que celles obtenues avec les données d'agriculteurs. Par exemple, pour atteindre une variance d'estimations égale à 4, il faut environ 8 données d'essais contre 230 données d'agriculteurs (figure 2).

L'analyse de la variance du modèle des données agrégées indique que 86 % de la variabilité du rendement observée entre les différentes variétés de blé est

### CONSTITUTION DES EFFECTIFS : six fois plus de données « agriculteur » que de données expérimentales



Figure 1

**Provenance des données d'essais et des données « agriculteur ».** Source : Arvalis (carte) ; Arvalis, Wiuz, Terrena & Axérial (données).

### PRÉCISIONS DES DONNÉES : il faut plus de données « agriculteur » pour atteindre une même variance

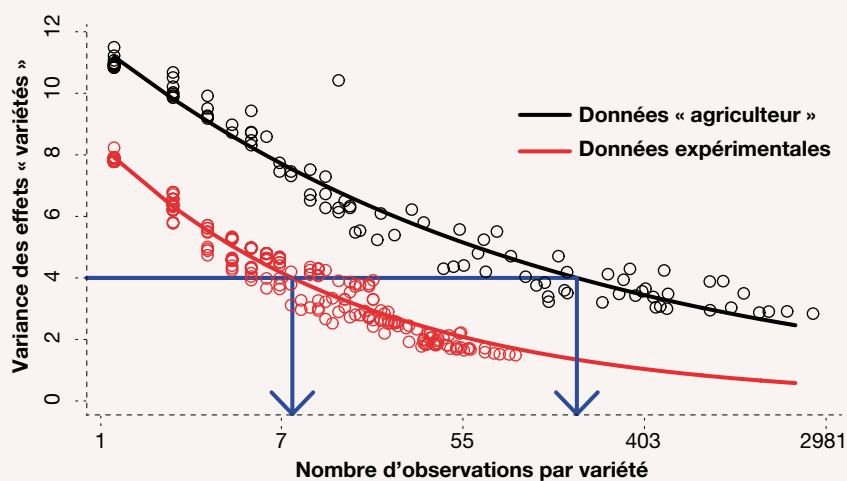


Figure 2

**Variance des effets « variété » en fonction du nombre d'observations par variété.** En rouge : données d'essais ; en noir : données « agriculteur » (échelle logarithmique). La variance diminue avec le nombre de données analysées ; plus la variance est faible, plus les estimations sont précises. À nombre de données égales, la courbe des données « agriculteur » est toujours au-dessus de la courbe des données d'essais, preuve que les données « agriculteur » sont plus hétérogènes et qu'il en faut donc plus pour obtenir une précision donnée.

commune aux deux types de données. Autrement dit, le niveau de cohérence entre les deux jeux de données vis-à-vis du classement des variétés selon leur productivité est de 86 %. L'estimation du rendement des variétés obtenue avec les données « agriculteur » est donc cohérente avec celle issue des données d'expérimentation.

Il existe cependant une interaction modérée entre les variétés et le type de données, représentant 14 % de la variabilité globale des classements variétaux. Les données disponibles n'ont pas permis d'expliquer plus précisément l'origine de cette interaction.

### Les données des deux réseaux apparaissent complémentaires

Les réseaux d'expérimentation classique fournissent des données de grande qualité, notamment sur les variétés récentes non encore largement diffusées chez les agriculteurs. Toutefois, elles sont obtenues pour des pédoclimats et des conduites de culture moins diversifiées.

Les données agriculteurs, quant à elles,

sont plus difficiles à exploiter car de nombreux facteurs varient entre parcelles et entre exploitations. En revanche, elles couvrent une gamme d'environnements plus large que les réseaux d'expérimentation et correspondent aux pratiques réelles des agriculteurs. De plus, ces données apportent de l'information sur les variétés plus anciennes, qui ne sont plus testées dans les réseaux expérimentaux.

La valorisation de données obtenues à d'autres fins que des fins de recherche, par exemple les données d'agriculteurs stockées dans les bases de données des logiciels de gestion parcellaire, semble être une tendance de fond. Et même si l'expérimentation agronomique reste un outil indispensable dans la production de connaissances, la question qu'il faut se poser n'est peut-être pas s'il faut essayer de valoriser ce type de données, mais comment les valoriser de manière optimale et rigoureuse.

La suite de ces travaux pourrait dès lors consister à déterminer comment combiner données « agriculteur » et données d'expérimentation, sans doute en leur

accordant un poids fonction de leur précision respective, afin de répondre au mieux aux questions qui nous sont posées. L'étape suivante serait de réfléchir à la façon d'optimiser les dispositifs expérimentaux dans l'optique de combiner les résultats issus de ces différents réseaux. ■

#### TÉMOIGNAGE

### UN GISEMENT ENCORE SOUS-EXPLOITÉ

“ Nous avons été particulièrement intéressés par la validation de la cohérence globale ainsi que par le niveau de précision obtenue. Les données « agriculteur » sont fiables : elles permettent d'établir un classement cohérent avec le classement obtenu depuis le réseau expérimental. Et elles sont relativement précises : 230 données « agriculteur » apportent la même précision que 8 données essais. Ce volume de donnée est finalement relativement faible au regard de l'utilisation croissante des outils (plus de 2500 agriculteurs utilisateurs chez Terrena). Nous n'avons pas donné suite à ces travaux mais restons persuadé qu'ils peuvent nous aider à améliorer notre conseil auprès des exploitants. Une telle valorisation de leurs propres données les conforterait également dans l'utilisation des outils numériques. À poursuivre donc... ”

ÉDOUARD BERTHELIN

Responsable équipe Data & Décisions  
chez Terrena

#### TÉMOIGNAGE

### UNE PLATEFORME QUI MUTUALISE LES DONNÉES NUMÉRIQUES DES AGRICULTEURS

“ L'usage numérique se développe très rapidement au sein de la plateforme Wiuz<sup>(1)</sup> : le nombre de parcelles numérisées y croît de 20 % par an. Ce gisement de données ouvre de nouvelles perspectives mais encore faut-il des acteurs pour s'en emparer. Car des masses aussi importantes de données suivies dans des conditions réelles exigent des compétences et des modèles de traitement particuliers. Cette expérience originale et réussie d'analyse de données « Variétés » à grande échelle sur des parcelles pluriannuelles est riche d'enseignement et de possibilités. Elle conforte le lien entre expérimentation et données au champ. Cela permet d'envisager que, progressivement, ces grandes quantités de données deviennent elles-mêmes une source majeure de connaissances et d'enseignements. ”

(1) Plus d'informations sur <https://wiuz.fr>

DAVID TALEC

Responsable partenaires

MARC HOPPENOT

PDG de WIUZ



Céréales

► Blé tendre d'hiver

# Gerry

La qualité pour tous



COTATION  
CTPS

**102.6%**

en % des témoins CTPS Sud  
de 2018/19\*

- Recommandé par la meunerie française
- Précoce et très productif
- Très résistant à la verse
- PS et protéines élevés

Retrouvez plus d'informations sur notre site internet :  
[www.agriobtentions.fr](http://www.agriobtentions.fr)

