

Rapport de stage de DUT Statistique et Informatique Décisionnelle
2012-2013

**Collecte et analyses statistiques de données
issues d'expérimentations sur le recyclage
agricole de produits résiduaux organiques dans
différents contextes agropédoclimatiques**

Pierre DESMET

Remerciements

Je tiens à remercier toutes les personnes qui ont participé de différentes façons à la réussite de ce stage.

Je remercie tout d'abord mes deux tutrices de stage, Alix Bell et Aurélia Michaud, pour leur attention et leur disponibilité tout au long de mon stage. Je les remercie également pour leur implication dans le présent rapport.

Mes amitiés vont à ma collègue et amie Laura Freudreich, pour son enthousiasme et sa bonne humeur, travailler à ses côtés a été un réel plaisir. Je la remercie d'ailleurs d'avoir répondu à (presque !) toutes mes questions concernant le domaine agricole, jusqu'alors inconnu.

Je tiens à remercier Vincent Mercier, pour m'avoir fait visiter l'essai au champs Qualiagro.

Je remercie Jean-Yves Cahurel et Xavier Delpuech pour m'avoir fourni rapidement les données sur lesquelles j'ai travaillé.

Merci à M. Fabien Couprie pour l'aide qu'il m'a apporté sur le logiciel Statbox par téléphone puis par mail.

Merci à Michel Levy de m'avoir suivi tout au long de mon stage et pour ses encouragements.

Je n'oublie pas de remercier tous mes enseignants du département STID de l'IUT Paris Descartes, en particulier Mme Florence Muri, M. Jérôme Dedecker et M. Olivier Bouaziz, pour leurs cours qui m'ont été très utiles.

Sommaire

Introduction	7
Présentation de l'entreprise	9
1 L'expérimentation agronomique de plein champ	11
1.1 Principe général de l'expérimentation en agronomie	11
1.2 Élaboration du plan d'expérience	12
1.2.1 Mise en place d'un essai expérimental au champ : définitions préalables	12
1.2.2 Mise en place d'un essai expérimental au champ : les dispositifs expérimentaux	13
1.2.3 Dispositif en blocs aléatoires complets	14
1.2.4 Dispositif en carré latin	15
1.3 Démarche statistique générale appliquée aux essais agronomiques	16
1.3.1 Validation agronomique et statistique d'un essai	16
1.3.2 Mise en évidence de différences entre les traitements testés	16
1.3.3 Précision et puissance	17
1.3.4 Ecueils de l'inférence statistique	17
2 Méthodologie adoptée et outils utilisés	19
2.1 Présentation des essais à saisir et méthode de travail	19
2.2 Saisie des données mesurées sur les essais	21
2.2.1 Fichier de saisie	21
2.2.2 Méthode de saisie	21
2.3 Logiciel et méthode statistique	22
2.3.1 Logiciel utilisé	22
2.3.2 Méthodes statistiques définies pour le réseau PRO : tests utilisés	23
3 Analyse statistique des jeux de données	30
3.1 Vérifications préliminaires	30
3.1.1 Cartographie des résidus	32
3.1.2 Égalité des variances dans les traitements et dans les blocs	33
3.1.3 Normalité des résidus	34
3.2 Test global de l'effet des facteurs traitement et bloc	34
3.3 Comparaison entre les traitements	35
3.4 Problèmes rencontrés et solutions statistiques	37
3.4.1 Présence de résidus suspects	37
3.4.2 Non normalité	38
3.4.3 Hétéroscedasticité	41
Conclusion	42
Annexes	45

Table des figures

1	L'expérimentation agronomique selon Letourmy	11
2	Exemple d'un dispositif en blocs aléatoires complets	15
3	Exemple d'un dispositif en carré latin	15
4	Plan du dispositif de l'essai Saint-Etienne	20
5	Méthode de validation d'un jeu de données pour un paramètre choisi, et mise en évidence de différences entre traitements pour ce paramètre	23
6	Exemple de cartographie Statbox	26
7	Sorties Statbox de statistiques descriptives et boxplot pour le rendement moyen en raisin (exprimé en kg/cep)	31
8	Histogramme réalisé sur Statbox	32
9	Cartographie des résidus de l'essai Saint-Etienne au 13 Septembre 2010	33
10	Histogramme des résidus de l'essai Saint-Etienne	34
11	Histogramme des résidus du poids moyen BDT non normal au 9 mars 2009	38
12	Sortie Statbox du test de Kruskal-Wallis	41

Liste des tableaux

1	Table générale d'analyse de la variance à deux facteurs dans le modèle sans interaction .	28
2	Données du rendement moyen par cep en raisin pour l'essai Saint-Etienne, estimation le 13 septembre 2010	30
3	Test de Bartlett de l'homogénéité des variances dans les traitements et dans les blocs . .	33
4	Table Statbox d'analyse de la variance à deux facteurs	35
5	Moyennes du rendement moyen en raisin obtenues pour chacun des traitement de l'essai St-Etienne	35
6	Plus petite amplitude significative pour 2, 3, 4 et 5 moyennes comparées	36
7	Test de Newman-Keuls de distinction de groupes homogènes parmi 5 traitements	36
8	Test de Newman-Keuls de comparaison de traitements	37
9	Les différentes transformations de variable possibles et les justifications correspondantes	38
10	Présentation des mesures du paramètre P total HF par traitement au 10 février 2009 . .	39
11	Tableau des étapes de calcul de la statistique H du test de Kruskal-Wallis	40
12	Table de Newman-Keuls donnant la valeur de t_s	49

Liste des sigles et abréviations utilisés

ANOVA : ANalysis Of VAriance

Arg : argile

Ca : calcium

CDV : compost de déchets verts

Corg : carbone organique

Cu : cuivre

CTO : composés traces organiques

EGC : Environnement et Grandes Cultures

ETM : éléments traces métalliques

Fe : fer

FB : fumier de bovins

Ha : hectare

IFV : Institut Français de la Vigne

INRA : Institut National de Recherche en Agronomie

ISB : indice de stabilité biologique

ITK : itinéraire technique

K : potassium

Kg : kilogramme

Mg : magnésium

MO : matière organique

MS : matière sèche

N : azote

OMR : ordures ménagères résiduelles

P : phosphore

Ph : potentiel hydrogène

PRO : Produit Résiduaire Organique

UMR : Unité Mixte de Recherche

Introduction

Chaque année, ce sont près de 330 millions de tonnes de produits résiduels organiques (PRO : effluents d'élevage, boues de station d'épuration, composts...) qui sont épandus en agriculture pour leur valeur agronomique (i.e valeurs amendante et fertilisante). Cependant, une gestion mal adaptée de ces produits peut présenter certains risques environnementaux (fuite de nitrates, émissions de gaz à effet de serre, etc.) et sanitaires (pathogènes). Il est donc important de bien connaître les caractéristiques de ces PRO et leurs possibles impacts sur le sol, les cultures, l'eau et l'air. De nombreux essais au champ visant à caractériser ces différents effets ont été mis en place en France par différentes structures, chacun présentant des méthodes différentes (Instituts de recherche, Chambres d'agriculture, etc.). Un des enjeux du projet Réseau PRO est de mettre en place la coordination en réseau des essais de plein champ étudiant les différents effets du recyclage des PRO.

Débuté en 2011, ce projet est piloté par l'Association de Coordination Technique Agricole (ACTA) ainsi que par l'équipe Environnement et Grandes Cultures SOL (EGC) de l'Unité Mixte de Recherche à l'Institut National de la Recherche Agronomique (INRA). Un des premiers objectifs de ce projet était de connaître les essais existants étudiant les différents effets de l'épandage des PRO. Une enquête nationale lancée en 2011 a ainsi permis d'inventorier 437 essais agronomiques, mis en place entre 1974 et 2012. Afin de mutualiser et de partager les données acquises sur l'ensemble de ces essais, celles-ci seront stockées dans un système d'information national.

Parmi les essais recensés lors de l'enquête figure un réseau d'essais conduits en viticulture, nommé Réseau MO (matière organique) et géré par l'Institut Français de la Vigne et du Vin (IFV). Ce réseau a pour objectif l'amélioration du conseil aux viticulteurs en termes de gestion de la matière organique du sol. Il comporte 7 essais, dans lesquels on teste les effets de PRO sur les compartiments plante, sol et PRO. Les données acquises sur les essais de ce réseau sont destinées à être insérées dans la base de données du Réseau PRO.

L'objectif de mon stage était dans un premier temps de recueillir, de mutualiser et de mettre en forme les données relatives à des essais du réseau MO. Les données sont mises en forme dans un fichier de saisie prédéfini, élaboré dans le cadre du Réseau PRO.

Afin d'assurer une transparence sur la qualité des données acquises sur les essais recensés dans le Réseau PRO, des critères statistiques et agronomiques seront associés aux données insérées en base. Une validation agronomique et statistique des données est donc nécessaire avant l'insertion des données en base.

Le travail de saisie m'a permis dans un second temps de valider statistiquement un certain nombre de jeux de données acquises sur les différents compartiments suivis au champ (sol, plante, PRO). Cette validation statistique des données permet de vérifier que celles-ci sont exploitables et respectent les conditions nécessaires aux tests et modèles d'analyse de la variance (ANOVA). Ainsi seront mises à disposition du public des données dont les qualités statistiques et agronomiques sont identifiables.

L'enjeu était de pouvoir valider les essais sur lesquels j'ai travaillé.

Dans une première partie nous nous attacherons à décrire l'expérimentation agronomique de plein champ. Nous détaillerons la démarche à entreprendre lors d'une expérimentation, avant de présenter l'intérêt d'un dispositif expérimental ainsi que les principaux dispositifs utilisés en agronomie. Nous décrirons alors la démarche statistique générale appliquée aux essais en agronomie.

Une deuxième partie sera consacrée à la méthodologie adoptée, et aux outils utilisés pour mener à bien notre mission. Les essais dont les données ont été saisies seront présentés, puis le fichier de saisie en lui même et la méthode de saisie. Nous décrirons ensuite le logiciel de travail utilisé ainsi que les tests statistiques.

L'analyse statistique des données fera l'objet d'une troisième et dernière partie. Y figurent des vérifications préliminaires, nécessaires à la validité de l'essai, ainsi que la vérification des hypothèses de l'ANOVA et des tests. Enfin nous jugerons de l'existence de différences entre les traitements, et, le cas échéant, nous les distinguerons en plusieurs groupes.

Présentation de l'entreprise

A. L'INRA

L'Institut National de Recherche Agronomique (INRA) est un organisme de recherche publique, qui mobilise de nombreuses disciplines scientifiques telles que les sciences de la vie, les sciences des milieux et des procédés, l'ingénierie écologique ou encore les écotechnologies et les biotechnologies. L'INRA produit des connaissances scientifiques et accompagne l'innovation économique et sociale dans les domaines de l'alimentation, de l'agriculture et de l'environnement. Géré par le ministère chargé de la Recherche et par le ministère chargé de l'Agriculture, l'INRA est le premier organisme de recherche agronomique en Europe et est le deuxième dans le monde en nombre de publications en sciences agricoles et en sciences de la plante et de l'animal.

L'institut est composé de 8504 agents permanents (chercheurs, ingénieurs et techniciens), répartis dans 17 centres de recherche régionaux, 13 départements scientifiques et plus de 200 unités de recherche, en particulier l'Unité Mixte de Recherche Environnement et Grandes Cultures (UMR EGC), au sein de laquelle s'est déroulé mon stage.

B. L'Unité Mixte de Recherche Environnement et Grandes Cultures

Si cette unité est qualifiée de mixte, c'est parce qu'elle est placée sous la double tutelle de l'INRA et d'AgroParisTech. Elle entend contribuer à faire évoluer l'agriculture dans le sens d'une meilleure adéquation avec les attentes de la société.

Le personnel de l'unité est réparti en trois équipes scientifiques (sol, biosphère-atmosphère et plante), ainsi que dans une équipe de service communs (ateliers, documentation, secrétariat et informatique). L'UMR est constituée majoritairement de chercheurs et d'enseignants-chercheurs (39) spécialistes des systèmes de culture, du sol ou des végétaux. Elle compte également 18 ingénieurs et 30 techniciens et administratifs.

L'objectif de l'UMR est d'approfondir les connaissances sur le fonctionnement des agrosystèmes du Nord de l'Europe et ses interactions avec les facteurs environnementaux.

L'unité conduit ainsi de nombreuses études scientifiques sur les sols, l'interface végétation-atmosphère et sur le couvert végétal. Elle fait partie de collaborations scientifiques nationales, européennes et internationales, et est impliquée dans des formations de l'enseignement supérieur (Ecole d'ingénieurs AgroParisTech, master de l'Université Pierre et Marie Curie). Mon stage s'est déroulé dans l'équipe Sol de l'unité, présentée ci-dessous.

C. Équipe SOL

L'équipe Sol travaille sur des problématiques agro-environnementales, structurées en 2 axes d'études : conséquences environnementales de l'utilisation de PRO sur le sol, et dynamique des pesticides et évaluations des impacts sur la qualité des sols, des eaux et de l'air.

Les travaux de l'équipe contribuent à la construction et l'alimentation de bases de données sur les

PRO, au développement d'outils et de modèles simulant la dynamique du carbone, de l'azote et des pesticides dans le sol ainsi que leurs effets sur l'environnement.

L'équipe compte actuellement 15 scientifiques (chercheurs ou enseignants-chercheurs), 8 ingénieurs, 8 techniciens, 9 doctorants et 3 post-doctorants.

C'est dans ce cadre que s'est déroulé mon stage, au cours duquel j'ai réalisé les travaux présentés dans ce rapport.

Chapitre 1

L'expérimentation agronomique de plein champ

1.1 Principe général de l'expérimentation en agronomie

Réaliser une expérimentation, c'est provoquer un phénomène afin d'en étudier les conséquences. Il s'agit de tester des hypothèses posées a priori : on cherche à comparer les effets de différents traitements sur une population expérimentale, et en respectant le principe "toutes choses étant égales par ailleurs" (Vilain, 1999).

Letourmy (1999) propose le schéma suivant pour illustrer l'expérimentation agronomique.

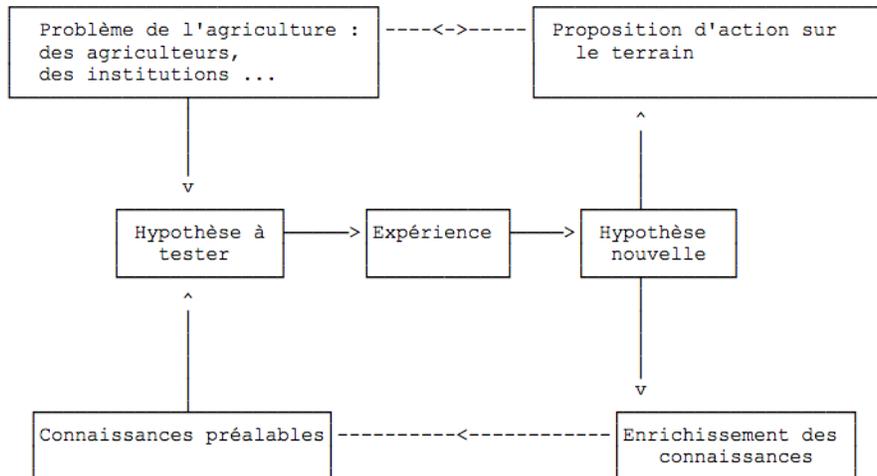


FIGURE 1 – L'expérimentation agronomique selon Letourmy

Certains problèmes ou questionnements peuvent se poser dans le monde de l'agriculture. Afin de répondre aux problématiques rencontrées par les agriculteurs, le rôle de l'expérimentateur en agronomie est de formuler une (ou des) hypothèse(s) pouvant expliquer les phénomènes auxquels on souhaite trouver des réponses. Ces hypothèses, formulées à l'aide des connaissances existantes, sont ensuite testées au moyen d'une expérience agronomique. Les résultats de l'expérimentation permettront de confirmer ou d'infirmer la (les) hypothèse(s) testée(s) voire de soulever d'autres hypothèses. L'expérimentation

au champ, lorsqu'elle est bien planifiée et confrontée aux connaissances préalables, est donc un outil d'aide à la décision.

La planification d'une expérimentation est essentielle et consiste, selon Vilain (1999), en 3 phases principales :

- l'élaboration du plan d'expérience,
- la réalisation et la conduite de l'essai,
- l'analyse et l'interprétation des résultats.

Ces trois phases seront explicitées en détail dans les paragraphes suivants.

1.2 Élaboration du plan d'expérience

1.2.1 Mise en place d'un essai expérimental au champ : définitions préalables

Le plan d'expérience est réfléchi en fonction de l'hypothèse que l'on souhaite tester. Cela nécessite de définir au préalable les facteurs pouvant influencer le paramètre que l'on étudie.

Un facteur est défini comme une série d'éléments de même nature (par exemple la nature d'un PRO, un fertilisant azoté, un engrais, une variété...) susceptibles d'influencer le résultat d'une expérimentation (par exemple le rendement de la culture fertilisée par un PRO donné). Il est défini par plusieurs variantes appelées modalités (facteur qualitatif) ou niveau (facteur quantitatif). On appelle facteurs étudiés les facteurs introduits volontairement par l'expérimentateur pour voir s'ils induisent des effets, et si oui lesquels. Ils sont l'objet même de l'expérience.

Prenons un exemple. On cherche à savoir s'il existe une différence de rendement en raisin lorsque l'on épand 2 PRO de natures différentes et à des doses d'apport différentes, par rapport au cas où l'on apporte aucun PRO : un fumier de bovins (noté FB), et un compost de déchets verts (noté CDV) apportés à 2 doses différentes raisonnées sur la dose d'azote apportée par les PRO épandus (60 kg N/ha et 80 kg N/ha) aux pieds de la vigne. Les facteurs étudiés sont donc :

- La nature du PRO, avec 2 modalités testées : FB et CDV
- La dose d'apport du PRO avec 2 niveaux : 60 kg N/ha et 80 kg N/ha

Lorsque l'on se place dans une expérimentation en plein champ, des facteurs extérieurs aux facteurs expérimentaux, testés sur le dispositif peuvent influencer le phénomène que l'on cherche à étudier. On appelle ces derniers des "facteurs aléatoires". Ils sont liés au milieu (ex. variabilité dans l'espace de propriétés du sol) ou au matériel expérimental (ex. changement de matériel en cours de travail sur le dispositif). Ces facteurs aléatoires, s'ils ne sont pas pris en compte, peuvent induire un biais expérimental sur les résultats acquis sur l'essai agronomique voire compromettre l'exploitation d'un essai. Certains facteurs aléatoires peuvent être pris en compte grâce à un dispositif expérimental adapté permettant de s'affranchir de leurs effets. Ces facteurs sont alors dits contrôlés (M. Vilain, 1999).

Lorsque l'on a identifié les facteurs à étudier ainsi que les variables suivies sur le dispositif expérimental (ex. rendement, teneur en azote, ...), on définit ensuite les traitements appliqués sur le dispositif. Un traitement est une combinaison de modalités/niveaux des facteurs étudiés. Afin de comparer les traitements à une situation de référence, il est courant d'inclure un traitement dit "témoin" pour lequel les modalités ou niveaux "zéros" sont attribués. Il n'est pas forcément nécessaire dans toutes les expérimentations (M. Vilain, 1999). Pour l'exemple choisi, les 5 traitements correspondant sont :

- Traitement T1 : FB × 60 kg N/ha
- Traitement T2 : FB × 80 kg N/ha
- Traitement T3 : CDV × 60 kg N/ha
- Traitement T4 : CDV × 80 kg N/ha
- Traitement « Témoin Zéro », sans apport de PRO ou d'azote

De manière générale, le nombre de traitements possibles correspond au produit du nombre de modalités/niveaux des facteurs étudiés, auxquels s'ajoute(nt) le(s) témoins définis.

Si un même traitement apparaît à plusieurs reprises, les parcelles qui le reçoivent sont considérées comme des répétitions de ce traitement. Les répétitions permettent d'avoir une mesure globale plus représentative d'un paramètre mesuré sur l'essai, d'intégrer la variabilité spatiale et expérimentale éventuellement rencontrée au sein du dispositif et d'effectuer des traitements statistiques sur les données acquises sur l'essai. Si aucune répétition n'est réalisée sur un dispositif (comme c'est le généralement le cas dans le cadre d'essais dits démonstratifs), la "robustesse" de l'essai ne sera pas suffisante pour permettre de réaliser des tests statistiques de comparaison entre traitements ni d'évaluer statistiquement l'effet global du/des facteurs testés. En effet, les valeurs obtenues de façon unique par traitement (y compris en cas de prélèvement multiples) si elles diffèrent ne pourront pas être mises en relation avec d'autres valeurs acquises dans les mêmes conditions expérimentales, c'est-à-dire pour les autres répétitions. Ainsi il est difficile d'assurer la "fiabilité" du résultat d'une valeur sans intégrer la variabilité expérimentale, celle-ci étant prise en compte dans le cadre des essais agronomiques à plusieurs répétitions. Si l'on ne faisait pas de répétition par exemple, le risque que la vraie valeur du paramètre diffère de celle que l'on mesure est de fait plus important. L'erreur est définie comme l'écart entre la vraie valeur d'un paramètre et la valeur de ce paramètre estimée, celle que l'on mesure.

Une fois que les traitements à étudier ont été définis, il est nécessaire de mettre en place le dispositif expérimental qui permettra de tester les effets des traitements sur le phénomène que l'on suit. Il existe plusieurs méthodes de construction de dispositifs expérimentaux. Ceux-ci obéissent à certains principes statistiques et permettent de s'adapter aux contraintes du terrain (variabilité connue, taille du dispositif, nombre de traitements, facteurs, itinéraire technique¹ le cas échéant). L'ensemble de ces notions seront abordées dans le paragraphe suivant.

1.2.2 Mise en place d'un essai expérimental au champ : les dispositifs expérimentaux

Afin de mettre en place une expérimentation agronomique au champ, il faut choisir une parcelle agricole sur laquelle se déroulera l'essai. Le choix de cette parcelle doit se faire de façon à ce qu'il y ait (i) le moins possible de facteurs aléatoires pouvant influencer les résultats de l'expérimentation et (ii) le plus d'homogénéité possible en particulier pour les variables suivies sur le dispositif ou pouvant influencer sur celles-ci. Il s'agit ensuite de définir le dispositif expérimental à mettre en place sur la parcelle agricole.

Il existe plusieurs types de dispositif expérimental (randomisation totale, blocs, carrés latins, split-plot etc.) que l'on choisit en fonction de trois critères : la nature et le nombre de traitements, la surface et l'homogénéité du terrain disponible, les contraintes liées au matériel et au personnel disponible.

Sur la parcelle agricole accueillant le dispositif expérimental choisi, chaque répétition de chaque traitement est affectée à une surface de terrain de taille constante appelée parcelle élémentaire ou unité expérimentale. La répartition dans l'espace de ces parcelles élémentaires dépend du type de dispositif choisi.

Deux principes fondamentaux sont à respecter lors d'une expérimentation :

- les répétitions : chaque traitement doit être affecté à plusieurs parcelles élémentaires, pour pouvoir prendre en compte la variabilité du milieu dans l'interprétation des tests ;
- la répartition aléatoire (randomisation) : les traitements doivent être répartis de façon aléatoire dans les parcelles élémentaires du dispositif expérimental afin d'intégrer aléatoirement l'éventuelle variabilité expérimentale. Cette randomisation est garantie par lancers d'une pièce de monnaie, tirage de papiers dans une urne, génération de nombres aléatoires etc.

1. L'itinéraire technique correspond aux interventions culturales effectuées sur l'ensemble du dispositif expérimental telle que la récolte, le travail du sol, l'apport de produits phytosanitaires, etc.

Ces 2 principes ont un intérêt statistique. Malgré les précautions que l'on aura pris soin de prendre dans le choix de la parcelle agricole, des hétérogénéités du terrain sont généralement rencontrées, par exemple une pente ou un gradient de fertilité (ex. gradient de teneur en azote ou en carbone, zones plus humides, ...) qui sont des hétérogénéités de terrain pouvant influencer sur le phénomène étudié et/ou les variables suivies et amplifier ce que l'on appelle l'erreur expérimentale.

L'erreur expérimentale est la somme d'une erreur de mesure du paramètre étudié et d'une erreur de randomisation elle-même engendrée par l'hétérogénéité entre les parcelles élémentaires du dispositif expérimental.

D'autre part, l'hétérogénéité diminue la précision des intervalles de confiance, et rend les estimateurs des paramètres moins précis. On répète les traitements pour avoir une idée de l'erreur expérimentale : sans répétition, on ne pourrait en effet pas l'estimer.

La répartition aléatoire des traitements au sein des parcelles élémentaires permet d'éviter une erreur systématique (biais) que causerait involontairement un expérimentateur en choisissant lui-même l'emplacement des traitements.

Le dispositif expérimental en randomisation totale est le dispositif qui répond le plus simplement à ces 2 principes : chaque répétition de chaque traitement est affectée de façon totalement aléatoire à une unité expérimentale (ou parcelle élémentaire). Cependant, ce type de dispositif peut s'avérer mal adapté aux possibles hétérogénéités existant sur le terrain ou encore aux contraintes matérielles et techniques, notamment en cas d'un nombre élevé de traitements.

Les dispositifs expérimentaux tels que les dispositifs en blocs ou en carré latin sont plus souvent rencontrés en expérimentation agronomique car mieux adaptés aux hétérogénéités du terrain. C'est pourquoi nous présenterons préférentiellement ces 2 types de dispositifs.

Il existe également des dispositifs qui permettent de s'affranchir de certaines contraintes techniques telles que les dispositifs en split-plot. Ces derniers ne seront pas développés ici car moins souvent rencontrés dans le cadre de l'expérimentation sur les PRO.

Pour étudier simplement l'effet des facteurs et leur interaction, il faut un plan d'expérience équilibré et orthogonal. Un plan d'expérience est dit équilibré si chacun des groupes défini par les traitements et par les blocs comporte le même nombre de parcelles. On appelle $n_{i\bullet}$ le nombre de parcelles qui reçoivent le traitement i , $n_{\bullet j}$ le nombre de parcelles affectées au bloc j , n_{ij} le nombre de parcelles qui vérifient ces deux conditions à la fois, et n le nombre de parcelles de l'essai. Deux facteurs sont qualifiés d'orthogonaux si la condition suivante est respectée :

$$\frac{n_{i\bullet}n_{\bullet j}}{n} = n_{ij}$$

1.2.3 Dispositif en blocs aléatoires complets

Dans le cas où un gradient existe sur la parcelle agricole de l'essai (variation de pente, de fertilité, proximité d'un cours d'eau...), il est possible de prendre en considération l'hétérogénéité connue en mettant en place un dispositif en blocs.

Un bloc est une partie de terrain supposée homogène, c'est-à-dire que l'on admet que le sol a la même constitution physique, chimique, topographique etc. Un bloc est constitué d'un ensemble de parcelles élémentaires recevant chacune une répétition de chacun des traitements étudiés. Il y a donc autant de blocs que de répétitions des traitements (voir figure 2). Afin de contourner le gradient observé sur la parcelle, les blocs doivent être disposés perpendiculairement à celui-ci. De cette manière, l'ensemble des répétitions de traitement issues d'un même bloc est affecté de la même manière par l'hétérogénéité. Les effets des traitements pourront ainsi être comparés entre eux sans que l'effet de l'hétérogénéité n'interfère. Pour chaque bloc, une répétition de chaque traitement est affectée aux parcelles élémentaires de façon aléatoire, via un tirage aléatoire sans remise (une fois un traitement placé, on ne souhaite pas tirer ce traitement à nouveau).

Un dispositif expérimental en blocs aléatoires complets a un triple avantage ; l'hétérogénéité et donc la variance de l'erreur expérimentale est mieux maîtrisée, les travaux sont plus faciles à exécuter, et il est facile de comparer visuellement les traitements puisqu'ils sont peu éloignés les uns des autres (M. Vilain).

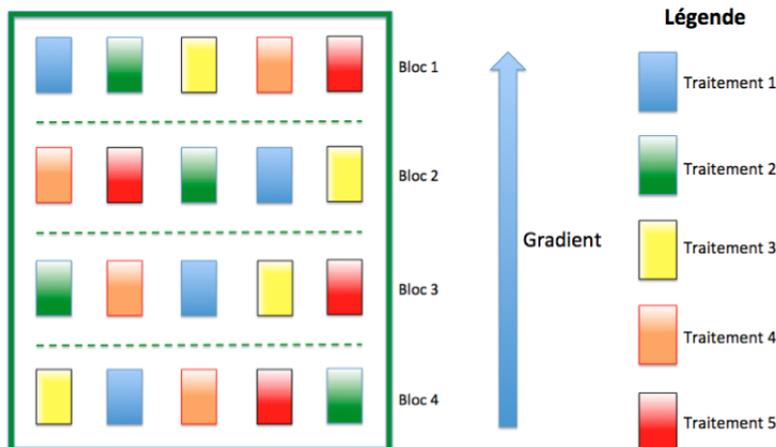


FIGURE 2 – Exemple d'un dispositif en blocs aléatoires complets

Chaque traitement de la figure 2 est affecté une et une seule fois dans chacun des blocs.

1.2.4 Dispositif en carré latin

Ce dispositif expérimental est souvent utilisé lorsque le nombre de traitements est compris entre 5 et 8, quand deux gradients sont perpendiculaires et si la direction du gradient est incertaine et peut traverser le champ en diagonale (M. Vilain). Un exemple de plan en carré latin est présenté en figure 3. Pour ce dispositif, il y a autant de traitements que de colonnes, de lignes et de répétitions d'un même traitement. Des blocs sont formés simultanément dans le sens des lignes et dans le sens des colonnes. Pour construire un dispositif en carré latin, on part d'un carré latin intuitif, de base. On permute les lignes aléatoirement, puis les colonnes.

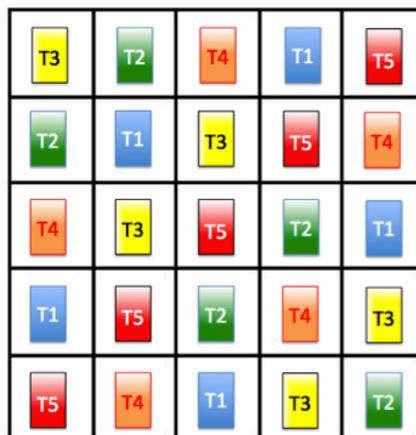


FIGURE 3 – Exemple d'un dispositif en carré latin

Sur la figure ci-dessus, les T_i représentent des traitements, et les rectangles des parcelles.

Après avoir présenté différents types de dispositifs expérimentaux et leur intérêt statistique, il faut valider l'essai. Cette validation se fait d'abord d'un point de vue agronomique, et ensuite seulement statistique.

1.3 Démarche statistique générale appliquée aux essais agronomiques

Nous allons présenter plus précisément les tests statistiques appliqués aux dispositifs en blocs aléatoires complets, puisque c'est sur ce type de dispositif que j'ai travaillé.

1.3.1 Validation agronomique et statistique d'un essai

Avant toute validation statistique d'un essai, il est d'abord nécessaire de valider la qualité agronomique de l'essai. Cette validation se fait via l'expertise de l'expérimentateur, ses connaissances agronomiques et son expérience. Il s'agit de vérifier que les conditions expérimentales n'ont pas été compromises de façon évidente par des facteurs non contrôlés tels que des aléas climatiques, l'invasion de ravageurs ou d'adventices sur la parcelle agricole, etc.

Dans le cas où l'expérimentateur juge que l'essai est acceptable d'un point de vue agronomique, on peut alors procéder à la validation statistique de l'essai, des jeux de données et à l'analyse statistique des valeurs mesurées.

L'interprétation des tests statistiques appliqués aux données doit toujours être associée à une interprétation agronomique. En effets, certains phénomènes ne trouvant pas d'explication statistique peuvent être attribués à une cause agronomique.

Pour valider un essai statistiquement, il faut vérifier que les conditions suivantes sont bien remplies :

- absence de valeur aberrante (outlier). Une donnée anormalement élevée ou faible peut mettre l'hypothèse de normalité en défaut, et augmente l'erreur expérimentale.
- absence d'interaction bloc \times traitement dans le cas des dispositifs en blocs

Si les conditions nécessaires à l'ANOVA sont vérifiées, l'expérimentateur passe à l'analyse statistique des données.

1.3.2 Mise en évidence de différences entre les traitements testés

Dans un essai, lorsque l'on étudie l'influence de différents traitements sur un paramètre donné, l'objectif est dans un premier temps de savoir s'il existe une différence significative des traitements entre eux. Afin de mettre en évidence ou non une telle différence, et après en avoir validé les conditions d'application, on effectue une analyse de variance (ANOVA) pour le paramètre suivi.

Dans le cas d'un dispositif en blocs, pour un paramètre mesuré donné, le modèle utilisé pour le test de l'ANOVA en l'absence d'interaction bloc \times traitement est le suivant :

$$\text{Valeur mesurée} = \text{Effet moyen} + \text{Effet traitement} + \text{Effet bloc} + \text{Effet résiduel}$$

Une valeur mesurée sur une répétition d'un traitement donné et dans un bloc donné, est influencée par l'effet moyen, c'est-à-dire l'effet de la parcelle agricole en général, par l'effet du bloc, par l'effet du traitement et enfin par un effet résiduel. L'effet résiduel traduit l'effet des facteurs non contrôlés, ou autrement dit, l'erreur expérimentale. Le résidu est donc l'écart observé pour chaque parcelle élémentaire entre la valeur réelle (observée) et la valeur théorique (valeur sans effet résiduel).

Dans le cas où le test de l'ANOVA permet de conclure à l'existence significative d'une différence entre les traitements, il s'agit ensuite de caractériser cette différence. En pratique, on utilise des tests paramétriques de comparaison de moyennes, tels que le test de Newman-Keuls, pour classer les traitements en groupes homogènes, ou le test de Dunnett, pour comparer les moyennes des traitements à un témoin.

1.3.3 Précision et puissance

Afin de connaître la pertinence des résultats des tests statistiques effectués, l'expérimentateur se réfère généralement à deux indicateurs statistiques : la précision de l'essai et la puissance de l'essai.

Pour connaître la **précision** d'un essai pour une variable donnée, on regarde l'écart-type résiduel de l'essai (ETR). Cet indicateur correspond à la somme de tous les résidus² de l'essai : c'est une estimation globale de l'erreur. Si l'on appelle ϵ_{ij} le résidu d'une mesure associée à un traitement i et prélevée dans le j -ème bloc, p le nombre de traitements de l'essai et q le nombre de blocs, on a donc la formule suivante :

$$ETR = \sum_{i=1}^p \sum_{j=1}^q \epsilon_{ij}$$

Exprimé dans la même unité que la variable analysée, il traduit la précision des résultats. La précision d'un essai est relative à l'ordre de grandeur des ETR généralement observés pour la même variable et dans des essais analogues. Généralement, un essai sera considéré comme précis si son ETR est inférieur ou égal à l'ETR habituellement observé pour des essais similaires. Pratiquement, c'est à chaque expérimentateur de se situer par rapport aux résultats habituels qu'il obtient selon l'année, l'espace étudié, la région naturelle, etc. (J.P. Gouet et G. Philippeau, 2002). Un ETR anormalement élevé peut provenir d'une hétérogénéité du terrain ou d'un excès de traitements expérimentés par exemple. Comme dans tout dispositif, le nombre de répétitions est fonction croissante de la précision des résultats de l'essai. Autrement dit, plus l'essai comportera un nombre élevé de répétitions, plus l'on pourra faire confiance aux résultats trouvés. Toutefois, accroître la taille de l'essai augmente le risque d'un biais expérimental dû à l'hétérogénéité du terrain.

La **puissance** d'un essai agronomique correspond à la probabilité de mettre en évidence une différence donnée entre traitements. Dans notre cas, elle révèle la capacité qu'aurait l'essai à mettre en évidence des différences entre traitements s'il y en avait. Elle est le complément à l'unité du risque β de deuxième espèce, soit $1 - P(\text{"ne pas rejeter l'hypothèse nulle alors que celle-ci est fautive"})$.

1.3.4 Ecueils de l'inférence statistique

Lorsque l'on observe les mesures d'un paramètre (par ex. la teneur en carbone organique lors d'un prélèvement du sol), on peut commettre 3 types d'erreurs.

- L'erreur de mesure, ou erreur technique, décrit une situation où la valeur obtenue pour une unité diffère de la vraie valeur. Cela peut être dû à l'instrument de mesure par exemple. Si cette erreur s'avère être trop grande, les résultats de l'expérimentation ne seront pas interprétables.
- L'erreur de randomisation est l'erreur causée par l'hétérogénéité des parcelles élémentaires. Il s'agit d'une variable aléatoire dont on peut étudier la distribution.
- L'erreur d'échantillonnage est l'erreur liée à l'aléa de la sélection d'un échantillon dans la population ou à la variabilité due à l'échantillonnage (ex. plusieurs préleveurs) : on ne peut pas mesurer toutes les molécules de carbone organique du sol sur une parcelle par exemple, alors on en prélève un échantillon que l'on espère représentatif. La valeur obtenue pour l'échantillon

2. Un résidu se définit comme la différence entre la valeur prédite par le modèle linéaire et la valeur que l'on observe effectivement.

(généralement dit composite, c'est-à-dire correspondant au mélange de plusieurs prélèvements élémentaires par parcelle élémentaire) est ensuite extrapolée à l'échelle de l'unité expérimentale ou parcelle élémentaire.

L'erreur expérimentale est définie comme la somme de ces trois erreurs.

Chapitre 2

Méthodologie adoptée et outils utilisés

L'objectif de mon stage était de mutualiser et de valider des jeux de données acquis sur des essais agronomiques en viticulture. Les essais sur lesquels portaient mon travail et la méthode de travail employée sont présentés dans les paragraphes suivants.

2.1 Présentation des essais à saisir et méthode de travail

Les essais exploités au cours du stage sont intégrés au réseau MO, qui est un réseau de 7 essais agronomiques conduits en viticulture et coordonnés par l'IFV (Institut Français de la Vigne et du vin). Les essais du Réseau MO ont été mis en place dans l'optique de mieux connaître les PRO épandus sur les parcelles viticoles et leurs effets sur le sol et la vigne, en particulier l'effet lié à l'apport de matière organique, connu pour jouer un rôle important dans le fonctionnement global du sol : (i) fonctionnement chimique puisque les matières organiques sont minéralisées par les microorganismes du sol en éléments minéraux assimilables par les plantes, et (ii) fonctionnement physique en améliorant notamment la structure du sol et les propriétés hydriques du sol. Les résultats acquis sur ce réseau ont pour but d'améliorer le conseil en termes de gestion de la matière organique auprès des viticulteurs. Mon travail a porté sur trois essais viticoles du réseau MO situés dans les villes d'Orgeuil (82), de Saint-Etienne-la-Varenne (69) et de Vauvert (30). Les caractéristiques de ces essais sont présentées dans des fiches descriptives en annexe.

Sur chacun de ces essais, un certain nombre de paramètres ont été mesurés pour chaque année d'expérimentation. Ces paramètres permettent de connaître la composition des PRO et de caractériser les effets de leur apport sur le sol et la vigne. Les mesures effectuées sur les PRO concernent leur caractéristiques physico-chimiques (teneur en matière sèche (MS), teneur en carbone organique (Corg), etc.). Pour caractériser les effets des PRO sur le sol des mesures ont été faites avant l'apport des PRO pour connaître l'état initial du sol, puis après l'apport des PRO. Ce sont, entre autres, des mesures physico-chimiques (pH, teneur en azote total (N total), teneur en Corg...). Pour les vignes, les paramètres suivis concernent les fruits (poids des baies, poids moyen de fruits par grappe...), le bois de taille (BT : c'est le bois obtenu lors de l'opération de la taille), les feuilles et les rognages (teneur en carbone (C) et en azote (N)...) et les pétioles (teneur en potassium (K), magnésium (Mg)).

Nous nous intéresserons ici plus particulièrement à l'essai se déroulant à Saint-Etienne-la-Varenne, pour lequel les résultats de la validation statistique des données seront présentés au chapitre 3 du rapport. Cet essai, débuté en 2009 et d'une durée de 8 ans, a pour objectif de connaître les effets de 2 stratégies de fertilisation, c'est à dire 2 PRO apportés à des fréquences différentes, sur les propriétés physico-chimiques du sol (teneur en matière organique du sol, teneurs en éléments fertilisants, etc.) et sur la vigne (rendement en raisins, teneurs en éléments fertilisants dans les différents organes de la vigne, etc.).

Cet essai est un dispositif en blocs aléatoires complets sur lequel est donc étudié le facteur stratégie de fertilisation organique, présentant 4 modalités. Dans le cas rencontré ici, puisqu'un seul facteur est étudié, les modalités correspondent à des traitements. Les 4 traitements organiques étudiés sont comparés à un traitement « témoin » où aucun apport de PRO n'est effectué. Les traitements appliqués sur le dispositif sont les suivants :

Témoin TEM : Traitement sans apport

Traitement ORG2 : apport d'Orgaveg (amendement organique du commerce) tous les 2 ans

Traitement ORG4 : apport d'Orgaveg tous les 4 ans

Traitement VEG2 : apport de compost végétal tous les 2 ans

Traitement VEG4 : apport de compost végétal tous les 4 ans

Les parcelles élémentaires sont identifiées par un numéro ou code. Le numéro de bloc est généralement le 1^{er} chiffre du numéro de la parcelle élémentaire, chiffre des centaines. Ensuite, les parcelles élémentaires à l'intérieur de chaque bloc sont notées de 1 à 10 ou plus en fonction du nombre de traitements par bloc.

La figure 4 ci-dessous présente le plan du dispositif.

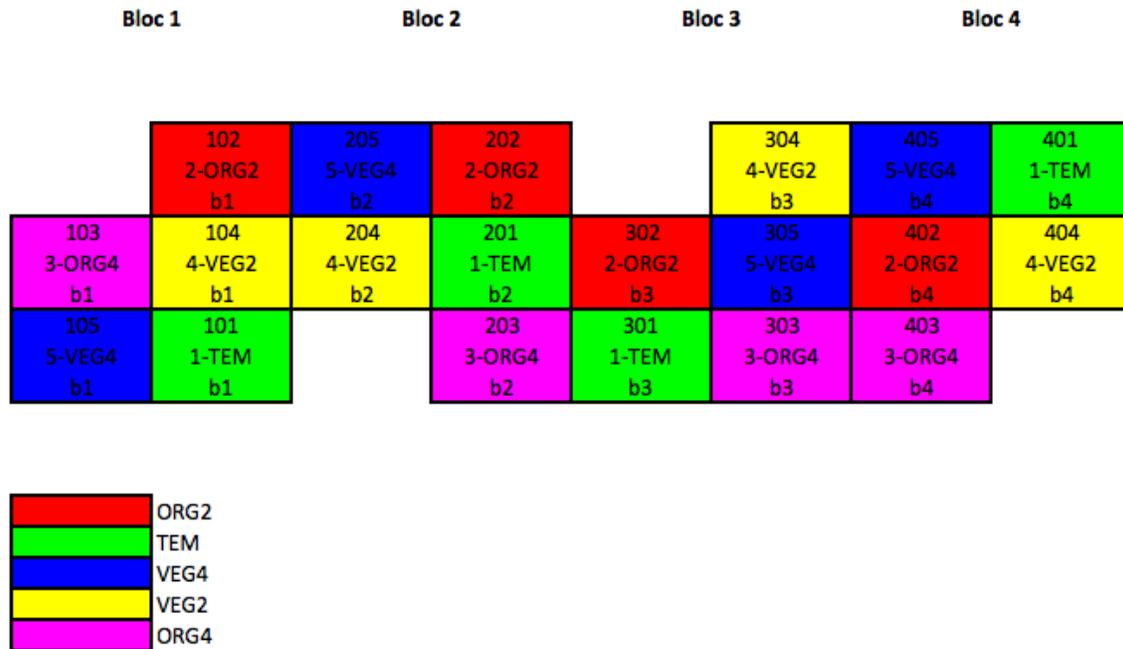


FIGURE 4 – Plan du dispositif de l'essai Saint-Etienne

Dans le cas de ce dispositif, les traitements appliqués comportent toujours le même numéro allant de 1 à 5. Par exemple, le traitement ORG2 portera toujours le numéro 2 quelque soit la localisation de la parcelle élémentaire dans le bloc. Ainsi, la parcelle 102 est située dans le bloc 1 et a reçu le traitement ORG2 par exemple (apport d'Orgaveg tous les 2 ans), la parcelle 202 est située dans le bloc 2 et a également reçu le traitement ORG2. Les termes b1 à b4 correspondent au bloc où figure la parcelle considérée.

2.2 Saisie des données mesurées sur les essais

Les valeurs mesurées pour chaque paramètre, chaque répétition de chaque traitement et pour chaque date de mesure ont été saisies dans un fichier Excel dédié, avec un fichier par essai. Le même fichier est rempli pour tous les essais. Celui-ci est présenté dans le paragraphe suivant.

2.2.1 Fichier de saisie

Afin de saisir et de mettre en forme toutes les données qui seront insérées dans la base de données du Réseau PRO dans un même format, un fichier Excel a été développé dans le Réseau PRO. Ce fichier permet de saisir dans différents onglets les informations relatives à l'essai nécessaires pour l'interprétation agronomique des résultats, telles que les objectifs de l'essai, l'historique sur 5 ans de la parcelle agricole, le contexte pédo-climatique, la description des PRO étudiés sur le dispositif, des traitements et de l'itinéraire technique.

Toutes les données relatives aux paramètres mesurés sur l'essai sont ensuite saisies dans 5 onglets spécifiques : mesures/analyses réalisées sur les PRO épandus sur le dispositif à chaque épandage, le sol récepteur des PRO, la plante cultivée, les reliquats azotés, les intrants (ex. phytosanitaires), l'eau et l'air. Dans ce fichier, pour un paramètre suivi donné, une valeur mesurée est systématiquement définie par sa date de mesure, le nom de la parcelle élémentaire, du traitement et du bloc sur laquelle elle a été mesurée ainsi que des informations associées au compartiment étudié (ex. la couche de sol considérée).

Le fichier d'insertion étant destiné à être repris pour insertion dans une base de données, il fallait que les informations renseignées pour chaque paramètre soient identiques sans erreur de saisie : cette homogénéité est rendue possible grâce à la présence de listes déroulantes, il y a un contrôle de saisie. L'interrogation de la base par des requêtes s'en trouvera ainsi grandement facilitée. En outre, à chaque valeur saisie est associé un statut qui permet d'ajouter une information qualitative sur la pertinence de la valeur, pour préciser par exemple si la valeur est "valeur mesurée", "valeur aberrante non remplacée", "valeur manquante", "valeur calculée", "valeur moyennée sur le traitement"...

2.2.2 Méthode de saisie

Après avoir pris contact avec le responsable du Réseau MO (Jean-Yves Cahurel), ce dernier m'a transféré l'ensemble des jeux de données des essais. J'ai complété les fichiers de saisie pour les 3 essais que j'ai pris en charge dans le cadre de mon stage, puis j'ai identifié un certain nombre d'informations manquantes, pourtant requises dans le fichier de saisie. Il manquait par exemple des informations sur le devenir des résidus de culture pour l'essai de Vauvert. J'ai donc pris contact avec les expérimentateurs des essais en question afin qu'ils me fournissent les informations complémentaires. Dans un second temps, afin de m'assurer de la qualité de ma saisie, j'ai contrôlé de façon aléatoire certaines valeurs saisies et les dates de mesure. Ma tutrice de stage a ensuite procédé de même.

Le fichier de saisie requiert également que les moyennes et les écart-types soient fournis. Ainsi, j'ai calculé pour chaque traitement et à chaque date de prélèvement les moyennes et écarts-types de chaque paramètre. Cette opération se justifie par la recherche d'un gain de temps pour le partage des données entre partenaires du Réseau PRO après leur mutualisation en base de données. En effet, la base de données finale sera complexe, et écrire un programme de calcul de ces agrégats demanderait plus de temps que de le faire manuellement dans le fichier de saisie. Les valeurs ainsi calculées ne sont pas les mêmes que celles mesurées par l'expérimentateur, elles bénéficient donc d'un statut différent "valeur moyennée sur le traitement".

Il faut aussi s'assurer que les données sont exprimées dans les mêmes unités. Cela se fait par simple inspection visuelle. Ce problème des unités différentes a été rencontré pour le paramètre *Poids Moyen Vendange*. Les informations relatives à ce paramètre provenaient de 2 sources différentes : l'une donnait le poids moyen de la récolte en raisin par mètre carré, et l'autre donnait la même information mais par pieds de vigne (cep).

Pour juger de la cohérence des données, on peut les comparer avec des valeurs attendues, celles que l'on observe généralement pour le paramètre mesuré. L'examen préalable de la cohérence des données est d'autant plus important que l'on utilise des moyens de traitement informatiques : si les valeurs d'un paramètre sont anormales d'un point de vue agronomique, ce n'est pas le logiciel qui nous le fera remarquer.

L'ensemble des données saisies dans le fichier ont ensuite fait l'objet d'une analyse statistique. Cette analyse permet de connaître la validité statistique de l'essai et des paramètres mesurés sur les essais insérés dans la base de données nationale du Réseau PRO. Elle permet également d'appliquer des analyses statistiques aux jeux de données en vue d'identifier d'éventuelles différences significatives entre traitements pour les variables mesurées. La méthode de saisie est présentée ci-dessous.

2.3 Logiciel et méthode statistique

2.3.1 Logiciel utilisé

Le logiciel Statbox PRO Agri est un module d'analyse et de traitement statistique intégré à Excel et développé par Arvalis – l'Institut du végétal. Le logiciel Statbox a été choisi pour les traitements statistiques des données du réseau PRO, puisqu'il a été développé spécifiquement pour les essais agronomiques et il présente des sorties et des résultats facilement interprétables et uniformes. Ce logiciel est ainsi abordable par le technicien comme par l'ingénieur. Bien que non rencontré dans ma formation, il m'a donc été facile de l'utiliser de part sa simplicité.

Le logiciel donne en quelques clics un ensemble de statistiques descriptives usuelles (moyennes, écart-type, quartile, étendue, Skewness, Kurtosis, coefficient de variation, boxplot...) pour une variable donnée. Lorsque l'on utilise Statbox pour tester des différences au niveau des paramètres selon les traitements, Statbox fournit en sortie :

- un test de l'interaction traitements \times blocs (test de Tukey) pour vérifier la validité du modèle retenu (pour certains dispositifs expérimentaux seulement)
- un histogramme des résidus pour vérifier leur normalité (approche qualitative) et les coefficients de symétrie et d'aplatissement de Pearson (mesure quantitative de l'écart par rapport à une loi de Gauss).
- les écarts-types des résidus intra traitements (et intra blocs) pour vérifier que les résidus ont la même dispersion dans tous les traitements (et dans tous les blocs).
- une cartographie des résidus sur le plan réel de l'essai pour vérifier l'indépendance des erreurs et identifier les résidus suspects sur le plan réel de l'essai.
- le tableau d'ANOVA qui permet de conserver ou de rejeter l'hypothèse d'homogénéité de l'ensemble des moyennes
- la puissance de l'essai pour apprécier les chances de réussite et aller plus loin dans l'interprétation des résultats dans le cas où l'effet traitement est non significatif.
- des analyses de variance sur les données, avec des tests de comparaisons multiples de moyennes (Newman Keuls).

2.3.2 Méthodes statistiques définies pour le réseau PRO : tests utilisés

Afin d'effectuer une validation statistique de l'essai et des jeux de données, puis d'effectuer les analyses statistiques pour identifier d'éventuels effets significatifs entre traitements, j'ai suivi de façon séquentielle et logique un protocole commun défini pour tous les essais agronomiques intégrés au Réseau PRO. Ce protocole est basé sur la méthode mise en œuvre par les statisticiens d'Arvalis. Celui-ci m'a guidé dans les interprétations et dans les décisions à prendre. Le logigramme suivant récapitule la démarche générale à avoir pour traiter statistiquement des données (figure 5).

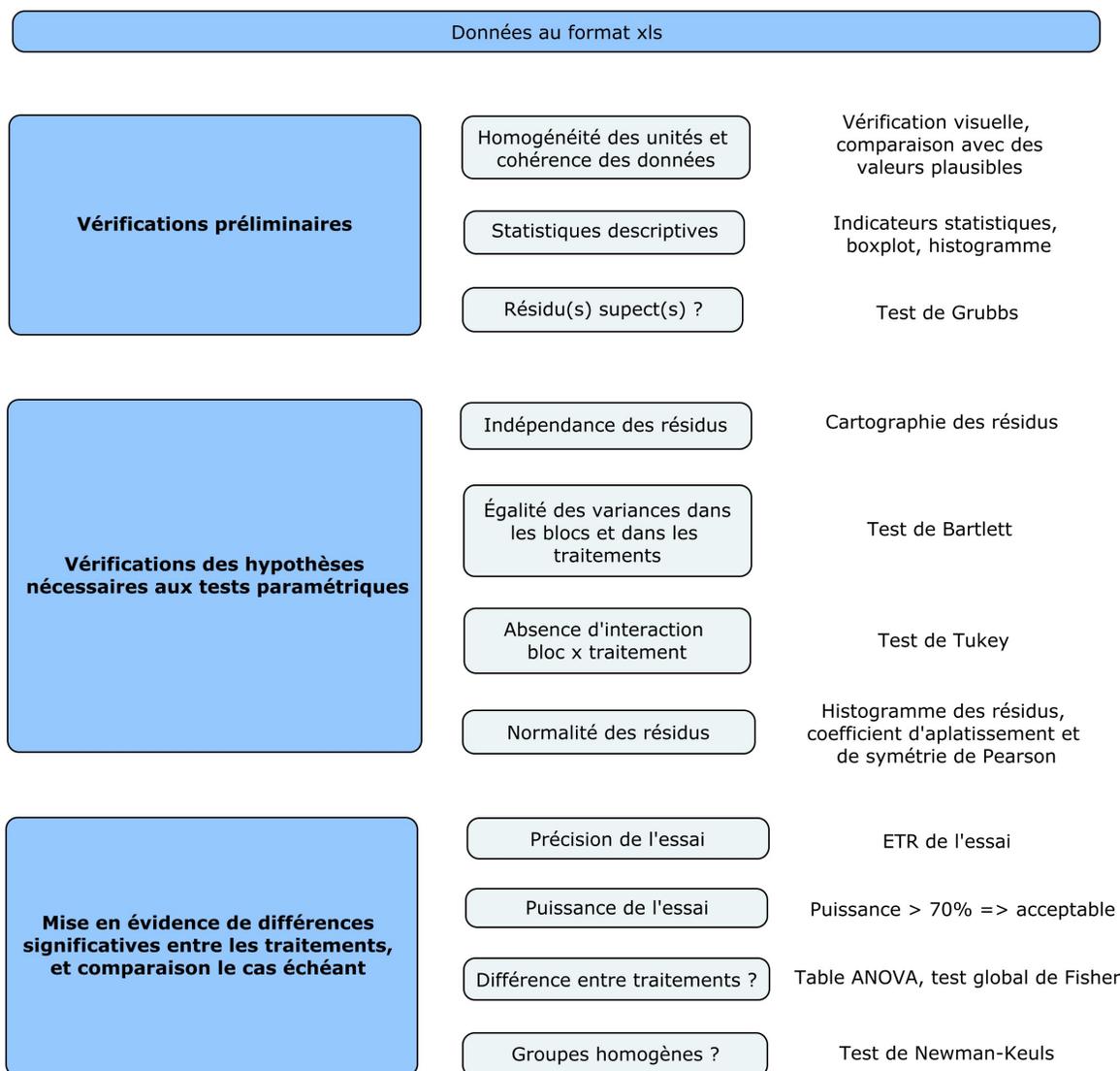


FIGURE 5 – Méthode de validation d'un jeu de données pour un paramètre choisi, et mise en évidence de différences entre traitements pour ce paramètre

L'ensemble des tests abordés par cette méthode sont explicités dans les paragraphes ci-dessous.

Vérifications préliminaires

Avant d'effectuer des tests, paramétriques ou non, sur un jeu de données, il convient de procéder à quelques vérifications.

Il est toujours recommandé de décrire un jeu de données statistiquement, à l'aide d'indicateurs et de graphiques. Des agrégats comme la moyenne ou la médiane renseignent sur la position d'un paramètre, et donnent une idée des grandeurs manipulées. Des graphiques comme les nuages de points ou les boîtes à moustache renseignent sur la dispersion d'un paramètre, et peuvent mettre en évidence d'éventuelles valeurs aberrantes (outliers), qu'un test de Tukey viendra peut-être corroborer. Le paragraphe suivant décrit les indicateurs statistiques les plus couramment utilisés.

La médiane est la valeur λ_2 de la distribution telle que la moitié des rendements moyens en raisin soient inférieurs à λ_2 . Si l'on désigne par y_k la k-ième mesure d'une variable continue et n le nombre de mesures d'un paramètre, la médiane Me est définie par la quantité suivante :

$$Me = \begin{cases} \frac{y_k + y_{k+1}}{2} & \text{si } n \text{ est pair} \\ y_{k+1} & \text{sinon} \end{cases}$$

Le premier quartile (respectivement le troisième quartile) est la valeur λ_1 (resp. λ_3) de la distribution telle que 75% (resp. 25%) des rendements moyens en raisin soient supérieurs à λ_3 .

Le coefficient de variation caractérise la dispersion de la distribution relativement à sa moyenne. On le calcule comme le rapport de l'écart-type de la distribution et sa moyenne. On considère souvent qu'au delà de 20% la variabilité de la distribution d'une variable par rapport à sa moyenne est élevée.

La moyenne géométrique est moins sensible que la moyenne arithmétique aux valeurs extrêmes d'une série de données. Sa formulation est la suivante :

$$\sqrt[n]{\prod_{k=1}^n y_k}$$

où les y_k sont les mesures positives du paramètre de l'essai, et n le nombre total de mesures.

Statbox détecte les valeurs anormales ou aberrantes, induites par des résidus suspects, en utilisant la méthode de Grubbs. En interne, le logiciel calcule \bar{y} , moyenne empirique d'une variable Y pour l'ensemble des observations y_i . L'écart-type empirique de la distribution s_i est aussi calculé. Pour chaque donnée y_i , Statbox calcule une statistique de test T_i , où

$$T_i = \frac{|y_i - \bar{y}|}{s_i}$$

Une observation est considérée comme anormale si sa statistique T_i associée dépasse un seuil T, la valeur de T figurant dans une table de Grubbs.

Analyse de Variance : vérification des conditions d'application du test

Lorsque l'on compare des traitements selon un dispositif en blocs aléatoires complets, le modèle statistique que l'on retient est le suivant :

$$Y_{ij} = \mu + \alpha_i + \beta_j + (\gamma_{ij}) + \epsilon_{ij}$$

Y_{ij} est la valeur observée d'un paramètre sur une parcelle de l'essai. Ladite parcelle a reçu le traitement i et figure dans le bloc j.

μ est l'effet moyen du champ d'essai, pour toutes les parcelles expérimentales, quelque soit le bloc et le traitement apporté.

α_i correspond à un effet du traitement i reçu par la parcelle.

β_j correspond à un effet du bloc j auquel appartient la parcelle.

γ_{ij} correspond à un effet d'interaction entre le i-ème traitement et le j-ème bloc. En agronomie, cet effet est souhaité nul : on travaille sur un modèle sans interaction, simplifié.

ϵ_{ij} correspond à une erreur résiduelle faite sur le i-ème traitement du bloc j. Le résidu est lié à l'aléatoire, à une possible erreur de mesure, à un éventuel mauvais choix du modèle, au terrain, ou au contexte pédo-climatique. Ces types d'erreurs sont malheureusement indissociables, seul le résidu "global" est calculable (J.P. Gouet, G. Philippeau).

Ce modèle est de dimension $p + q - 1 = 8$, où p et q représentent respectivement le nombre de traitements expérimentés et le nombre de blocs aléatoires complets de l'essais. On notera donc ce modèle M_8 . M_8 est additif, c'est-à-dire que la valeur observée du paramètre est le résultat de l'addition de plusieurs effets.

La valeur estimée par le modèle sans interaction est $\widehat{Y}_{ij} = \widehat{\mu} + \widehat{\alpha}_i + \widehat{\beta}_j$.

Ces paramètres sont respectivement estimés par les grandeurs suivantes :

$\widehat{\mu}$ est la moyenne empirique, calculée pour le paramètre sur toutes les parcelles expérimentales ;

$\widehat{\alpha}_i$ se calcule comme la différence entre la moyenne empirique du paramètre sur le traitement i et $\widehat{\mu}$;

$\widehat{\beta}_j$ s'obtient comme la différence entre la moyenne empirique des rendements du bloc j et $\widehat{\mu}$;

On peut calculer le résidu ϵ_{ij} pour chaque mesure :

$$\epsilon_{ij} = Y_{ij} - \widehat{Y}_{ij}$$

Certaines conditions propres à l'ANOVA et plus généralement aux tests paramétriques doivent être vérifiées avant l'analyse.

Les résidus calculés doivent tout d'abord être indépendants entre eux, c'est-à-dire que l'erreur commise sur une parcelle ne doit pas être liée à l'erreur faite sur une parcelle voisine. Statbox propose une manière visuelle de vérifier cette hypothèse : la cartographie des résidus. Cette représentation affiche la répartition des résidus sur le plan réel de l'essai. Chaque parcelle est colorée selon un gradient de couleur traduisant la valeur de son résidu. Cartographier les résidus permet de mettre en évidence un éventuel gradient intra-bloc, ou au contraire d'apprécier visuellement l'indépendance des résidus entre eux.

Cartographie des résidus :

	Bloc 1	Bloc 2	Bloc 3	Bloc 4
	1	2	3	4
1	Blue	Cyan	Blue	Blue
2	Cyan	White	White	White
3	Dark Blue	White	Cyan	Cyan
4	White	White	White	White
5	White	White	White	Blue
6	White	Dark Blue	Blue	Dark Blue
7	Dark Blue	Cyan	Dark Blue	Blue
8	White	Blue	Blue	Blue
9	Dark Blue	Dark Blue	Blue	Dark Blue
10	Dark Blue	Dark Blue	Dark Blue	Dark Blue

Légende :

White	Donnée manquante
White	< - 0,391
Cyan	< 0,000
Blue	< 0,391
Dark Blue	< 999999,000

FIGURE 6 – Exemple de cartographie Statbox

Égalité des variances inter-bloc et inter-traitement

Les résidus calculés doivent tous être du même ordre de grandeur, indépendamment du traitement ou du bloc : les résidus pour les différents traitements ou pour les différents blocs ne doivent pas être trop éloignés. Le logiciel teste l'homoscédasticité des résidus au moyen d'un test du Khi Deux. L'égalité des variances est une hypothèse raisonnable, les erreurs sont du même ordre de grandeur quelque soit le traitement expérimenté (J.P. Gouet et G.Philippeau, 2002). Le test de Bartlett teste l'homogénéité des écarts-types sur plusieurs populations (ici des traitements ou des blocs) au niveau de confiance $1 - \delta$, c'est-à-dire que l'expérimentation a $\delta\%$ de chance de fournir une conclusion erronée si des différences entre les variances ressortent. On teste ainsi

(H0) : *Les variances sont identiques pour tous les traitements (respectivement entre tous les blocs)* contre

(H1) : *Il existe au moins deux variances significativement différentes entre les traitements (respectivement dans tous les blocs)*

On calcule la statistique suivante :

$$T = \frac{[(\sum_{i=1}^p n_i) - p] \ln(S^2) - \sum_{i=1}^p (n_i - 1) \ln(s_i^2)}{1 + \frac{1}{3(p-1)} [\sum_{i=1}^p (\frac{1}{n_i-1}) - \sum_{i=1}^p \frac{1}{(n_i-1)}]}$$

n_i est le nombre de parcelles élémentaires recevant le traitement T_i (respectivement les parcelles élémentaires composant le bloc i , lorsque l'on cherche à tester l'égalité des variances pour les blocs) ;

p correspond au nombre de traitements (respectivement de blocs) ;

S^2 est l'estimateur groupé¹ (pooled) de la variance σ^2 du paramètre que l'on observe sur l'essai :

1. L'estimateur de la variance groupé est une méthode d'estimation de la variance globale de plusieurs échantillons, prélevés dans des circonstances différentes. Les moyennes peuvent varier entre les échantillons, mais la "vraie variance" est supposée être la même.

$$S^2 = \frac{1}{\sum_{i=1}^p (n_i - 1)} \sum_{i=1}^p (n_i - 1) s_i^2;$$

et s_i^2 est l'estimateur sans biais classique de la variance dans le traitement i (resp. le bloc i) : $s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$;

Y_{ij} est la mesure d'un paramètre de l'essai, mesuré dans le bloc j et ayant reçu le traitement i . La moyenne des mesures pour le traitement (ou le bloc) i est notée \bar{Y}_i .

La statistique de test T est d'autant plus élevée que les variances sont différentes. Sous (H_0) , on a la relation suivante : $T \sim \chi^2(p - 1)$, où p correspond au nombre de traitements (ou au nombre de blocs). Cette relation étant approximative, le test de Bartlett a de mauvaises performances sur les échantillons de petites tailles. L'hypothèse nulle est rarement rejetée du fait d'un test peu puissant (M. S Bartlett, 1937).

On rejette l'hypothèse nulle au risque δ si $T > \chi_{p-1;\delta}^2$, où $\chi_{p-1;\delta}^2$ est tel que $P(T > \chi_{p-1;\delta}^2) = \delta$.

Vérification de l'absence d'interaction entre l'effet traitement et l'effet bloc

Une interaction traitement \times bloc signifie que, pour un paramètre donné, les écarts de mesure entre les traitements ne sont pas les mêmes dans les différents blocs. Il existerait une interaction entre les caractéristiques du bloc et les traitements ayant un effet sur les valeurs mesurées. Dans ce cas, l'essai est biaisé et l'interprétation des résultats devient très délicate.

Pour tester la présence d'une éventuelle interaction multiplicative entre les traitements et les blocs, Statbox réalise un test d'additivité, ou test de Tukey.

Normalité des résidus

Une dernière hypothèse à vérifier moins importante est l'hypothèse que les résidus sont distribués normalement. Un moyen visuel de vérifier la normalité est de regarder l'allure de l'histogramme des résidus. S'il ressemble à une courbe en cloche, l'expérimentateur conclura à la normalité. Cette manière de procéder est un peu arbitraire : les conclusions sur un même histogramme peuvent différer d'un statisticien à l'autre, et l'histogramme peut être difficilement interprétable pour les petits essais. Afin de palier cela, Statbox propose le calcul de 2 indicateurs en complément, pour nous aider à mieux juger de la normalité de la distribution résiduelle.

- Le coefficient d'asymétrie de K. Pearson² (noté β_1) renseigne sur une éventuelle dissymétrie de la distribution d'une variable aléatoire réelle par rapport à sa moyenne. Sa forme explicite est la suivante :

$$\beta_1 = n \frac{[\sum_{k=1}^n (y_k - \bar{y})^3]^2}{[\sum_{k=1}^n (y_k - \bar{y})^2]^3}$$

où n est le nombre total de parcelles, y_k est la mesure du paramètre sur la k -ième parcelle, \bar{y} la moyenne générale du paramètre sur l'essai. Positif (resp. négatif), le coefficient d'asymétrie traduit une distribution étalée vers la droite (resp. vers la gauche). Une distribution parfaitement symétrique comme celle de la loi normale a un coefficient d'asymétrie nul. Le statisticien peut aussi comparer la moyenne de la distribution à la médiane : si la distribution est symétrique, ces deux valeurs seront proches.

2. Il ne faut pas confondre cet indicateur avec le coefficient d'asymétrie de Fisher (Skewness). Ces deux indicateurs s'interprètent de la même façon, mais le coefficient d'asymétrie de Pearson est le carré du coefficient d'asymétrie de Fisher.

- Le coefficient d'aplatissement de K. Pearson (noté β_2) renseigne sur le degré d'aplatissement d'une variable aléatoire réelle par rapport à une distribution gaussienne. Il est de la forme suivante :

$$\beta_2 = \frac{1}{n} \sum_{k=1}^n \left(\frac{y_k - \bar{y}}{\sigma_y} \right)^4$$

Où σ_y est l'écart-type de la distribution.

Une distribution parfaitement gaussienne a un coefficient d'aplatissement égal à 3. Plus cet indicateur est proche de 1, plus la distribution est aplatie (on parle de distribution platykurtique). A contrario, un coefficient d'aplatissement éloigné positivement de 3 est signe d'une distribution plus pointue que celle de la loi normale (distribution leptokurtique).

Analyse de variance

Une fois ces hypothèses vérifiées, nous voulons savoir si les traitements font varier le paramètre étudié de manière significativement différente. Répondre à cette problématique nécessite de réaliser une table d'ANOVA. Nous effectuerons un test global de Fisher dans lequel nous testerons au risque δ les hypothèses suivantes :

(H0) : *Les traitements ont le même effet sur le paramètre considéré* contre

(H1) : *Il existe au moins 2 traitements tels que les effets sur le paramètre considéré sont significativement différents.*

Le tableau d'analyse de la variance ci-dessous montre comment est calculée la statistique de test. Dans le cas d'une expérience faisant intervenir un seul facteur, l'interprétation des résultats la plus courante est basée sur l'analyse de la variance à deux critères de classification (P. Dagnelie, 1981), puisque le facteur subsidiaire bloc s'ajoute au facteur principal.

Il s'agit en fait d'un test global d'ANOVA à deux facteurs : le facteur étudié *Traitement*, fixe, et le facteur aléatoire *Bloc*, contrôlé. L'introduction du facteur *Bloc* n'a pour but que de rendre le test plus puissant, en vue de contrôler l'hétérogénéité et donc de réduire la somme des carrés résiduelle.

	DDL	Somme Des Carrés	Carrés Moyens	F
Variance Totale	$n - 1$	SCT	$SCT \div (n - 1)$	
Variance Traitement	$p - 1$	SCM_T	$SCM_T \div (p - 1)$	$\frac{SCM_T \div (p-1)}{SCR \div (n-p-q+1)}$
Variance Bloc	$q - 1$	SCM_B	$SCM_B \div (q - 1)$	$\frac{SCM_B \div (p-1)}{SCR \div (n-p-q+1)}$
Variance Résiduelle	$n - p - q + 1$	SCR	$SCR \div (n - p - q + 1)$	

TABLE 1 – Table générale d'analyse de la variance à deux facteurs dans le modèle sans interaction

Pour la Table 2.1 ci-dessus, nous détaillons les abréviations utilisées :

On désigne par n, n_i, n_j, p, q respectivement le nombre total de parcelles de l'essai, du traitement i , du bloc j , le nombre de traitements et le nombre de blocs.

$SCT = \sum_{k=1}^n (Y_k - \bar{Y})^2$ est la somme du carré des écarts de chaque point à la moyenne échantillonnée. Sous Excel, on peut retrouver cette valeur à l'aide de la fonction SOMME.CARRÉS.ÉCARTS().

$SCM_t = \sum_{i=1}^p n_i \bar{Y}_i^2 - n \bar{Y}^2$ est la somme des carrés expliqués par un modèle statistique où le traitement seul influence le paramètre. C'est une mesure de la variabilité de la moyenne du paramètre pour le traitement i \bar{Y}_i par rapport à la moyenne globale \bar{Y} .

$SCM_B = \sum_{j=1}^q n_j \bar{Y}_j^2 - n \bar{Y}^2$ est la somme des carrés expliquée par un modèle statistique où le bloc seul influence le paramètre.

$SCR = \sum_{i=1}^p (n_i - 1) s_i^2$ exprime la variabilité des observations Y_{ij} par rapport à la moyenne \bar{Y}_i de chaque traitement. s_i^2 est l'estimation de la variance du paramètre considéré pour le Traitement i . La variation totale du paramètre considéré se décompose alors ainsi :

$$SCT = SCM_T + SCM_B + SCR$$

La colonne Carrés Moyens (CM) divise les Sommes des Carrés (SC) par leurs degrés de libertés (DDL) respectifs. Les deux statistiques de test F s'obtiennent en divisant les CM par le carré moyen de la variance résiduelle. Il se trouve que ce rapport suit une loi de Fisher sous (H0).

$$\frac{SCM_T \div (p - 1)}{SCR \div (n - p - q + 1)} = F \sim F(p - 1; n - p - q + 1)$$

En interne, un logiciel compare la valeur de la statistique de test avec une valeur critique : on rejette l'hypothèse (H0) si $F > F_{p-1; n-p-q+1; \delta}$, avec $P(F > F_{p-1; n-p-q+1; \delta}) = \delta$. La valeur critique peut être lue dans une table de Fisher, mais tous les logiciels fournissent la p-valeur.

La p-valeur d'un test est le plus petit niveau pour lequel le test est significatif, c'est-à-dire tel que l'hypothèse nulle soit rejetée. Ici, elle correspond à la probabilité que la variable aléatoire F soit supérieure à sa réalisation f, celle que l'on observe. Statbox fourni également la p-valeur associée au test global de Fisher. La p-valeur est la probabilité d'observer une valeur de f aussi élevée.

Il est important de comprendre que l'absence de mise en évidence de différences entre les traitements ne prouve pas leur inexistence. Il est possible qu'une ou plusieurs moyennes diffèrent mais que leur différence n'ait pas pu être montrée dans les conditions opératoires présentes. Aucun moyen ne permet de savoir si un effet existe réellement, ou si les conditions n'ont pas permis de le mettre en évidence. (M. Vilain, 1999)

Si aucune différence significative entre les traitements n'a été révélée par l'ANOVA précédente, on arrête là l'analyse après avoir consulté la puissance et l'ETR de l'essai. Dans le cas contraire, on cherche à mettre en évidence des groupes de traitements homogènes, pour les classer les uns par rapport aux autres.

Détermination des différences existant entre les traitements

Le test paramétrique de Newman-Keuls permet de faire des comparaisons multiples de moyennes pour un paramètre, en comparant les moyennes 2 à 2. C'est ce test qui est retenu dans la méthode du réseau PRO. L'intérêt de ce test est qu'il ne nécessite pas la comparaison de tous les couples de traitements. En effet, le test de Newman-Keuls réduit le nombre total de comparaisons possibles³ en fonction des résultats des premières comparaisons pratiquées sur les moyennes classées. Le plan d'expérience doit être équilibré. Ce test étant assez complexe, on se propose de le voir plus en détail avec un exemple au chapitre 3 pour des raisons pratiques et de compréhension.

3. soit $C_n^2 = \frac{n!}{2(n-2)!}$ comparaisons possibles parmi n traitements...

Chapitre 3

Analyse statistique des jeux de données

L'ensemble de la démarche statistique employée sera illustrée par l'exemple de l'essai situé à Saint-Etienne-la-Varenne, présenté au chapitre 2.1.

3.1 Vérifications préliminaires

Dans l'ensemble du chapitre, on s'intéressera à la variable aléatoire réelle Y , obtenue sur l'essai de Saint-Etienne : elle correspond au rendement moyen en raisin, exprimé en kg/cep. On considère ce paramètre en particulier pour le rapport, parce qu'il est plus facile de l'appréhender lorsque l'on n'est pas agronome. Le rendement en raisin est mesuré de façon individuelle pour chaque parcelle élémentaire, ce qui correspond à 20 mesures du rendement moyen par cep en raisin (5 traitements répétés 4 fois). Pour rappel, le dispositif expérimental est en blocs aléatoires complets. Les traitements suivants sont répétés 4 fois sur le dispositif :

TEM : témoin sans apport organique

ORG2 : apport d'Orgaveg (amendement organique du commerce) tous les 2 ans

ORG4 : apport d'Orgaveg tous les 4 ans

VEG2 : apport de compost végétal tous les 2 ans

VEG4 : apport de compost végétal tous les 4 ans

Les données de rendements pour la vendange du 13 Septembre 2010 sont présentées dans le tableau suivant. À cette date, l'essai était déjà mis en place depuis un an. Le choix de cette période dans ce rapport se justifie par la mise en évidence de différence entre les traitements à cette date. Les résultats trouvés me permettent donc d'illustrer la démarche statistique du Réseau PRO de façon complète.

Bloc \ Traitement	TEM	ORG2	ORG4	VEG2	VEG4
Bloc 1	2,179	2,522	2,326	1,603	1,928
Bloc 2	1,679	2,489	2,436	1,883	2,160
Bloc 3	1,590	2,471	2,116	1,713	1,992
Bloc 4	1,014	2,986	2,576	2,251	1,935

TABLE 2 – Données du rendement moyen par cep en raisin pour l'essai Saint-Etienne, estimation le 13 septembre 2010

Dans notre exemple, le plan est équilibré, puisqu'il y a une unique mesure pour chaque couple Traitement \times Bloc. De ce fait, le plan est aussi orthogonal, n_{ij} est une constante égale à 1.

Conformément à la première partie du logigramme de la partie 2.3.2, j'ai vérifié que mes données étaient bien exprimées dans les mêmes unités, et que leur ordre de grandeur était cohérent par rapport à ceux des valeurs observées pour les années précédentes.

Comme pour chaque essai, j'ai calculé une série d'indicateurs statistiques, présentés ci-dessous.

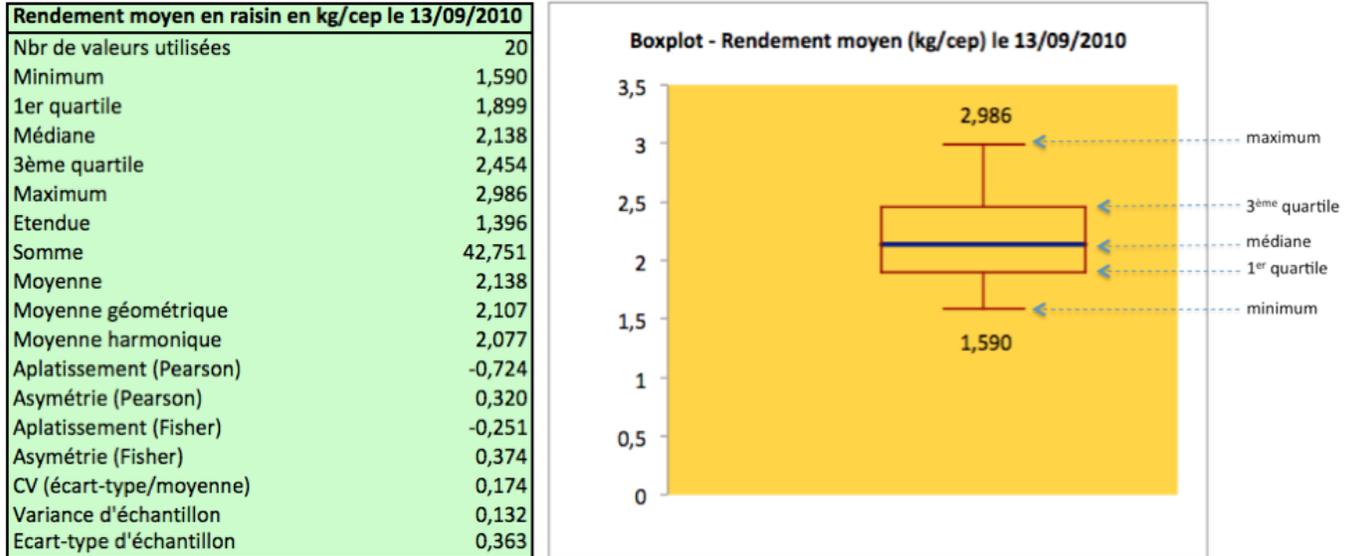


FIGURE 7 – Sorties Statbox de statistiques descriptives et boxplot pour le rendement moyen en raisin (exprimé en kg/cep)

La figure 7 présente des statistiques descriptives (à gauche) que l'on regarde dans un premier temps pour avoir une meilleure connaissance du jeu de données obtenu sur le dispositif et éventuellement mettre en évidence des anomalies (par exemple un écart-type anormalement élevé ou une distribution très asymétrique).

Elle nous apprend qu'un cep a en moyenne un rendement en raisin de 2,14 kg, quelque soit le traitement. Si la distribution possède des valeurs extrêmes influant sur la moyenne des résultats, il est préférable de s'intéresser à la médiane.

La médiane nous informe que 10 des 20 ceps ont un rendement moyen supérieur à 2,14 kg.

5 ceps ont un rendement moyen en raisin supérieur à 2,45 kg.

Le rendement moyen total de la récolte sur l'ensemble du dispositif s'élève à 42,75 kg de raisin.

Le cep le moins productif a un rendement moyen en raisin de 1,59 kg, tandis que le cep le plus productif produit en moyenne 2,986 kg de raisin.

La moyenne géométrique et arithmétique n'étant pas trop différentes, on peut se dire que les données sont exemptes de valeurs aberrantes.

Ces indicateurs nous donnent déjà une idée de la position et de la dispersion de nos données pour l'ensemble du dispositif. Il est possible de représenter certains d'entre eux graphiquement dans ce que l'on appelle une boîte à moustache (ou boxplot). Cela donne un aperçu plus visuel de la distribution de la variable.

La boîte à moustache présentée sur la figure 7 (à droite) ne met pas en évidence de valeur aberrante. La valeur de la médiane est d'ailleurs égale à celle de la moyenne. La boîte à moustache montre une distribution plutôt symétrique, et les valeurs des quartiles, relativement proches de la médiane (premier quartile de 1,9 et troisième quartile de 2,45) témoignent d'une dispersion faible du jeu de données.

D'autre part, une représentation graphique des valeurs mesurées par traitement permet d'avoir une première idée de la tendance de l'effet des traitements sur le paramètre suivi (voir figure 8).

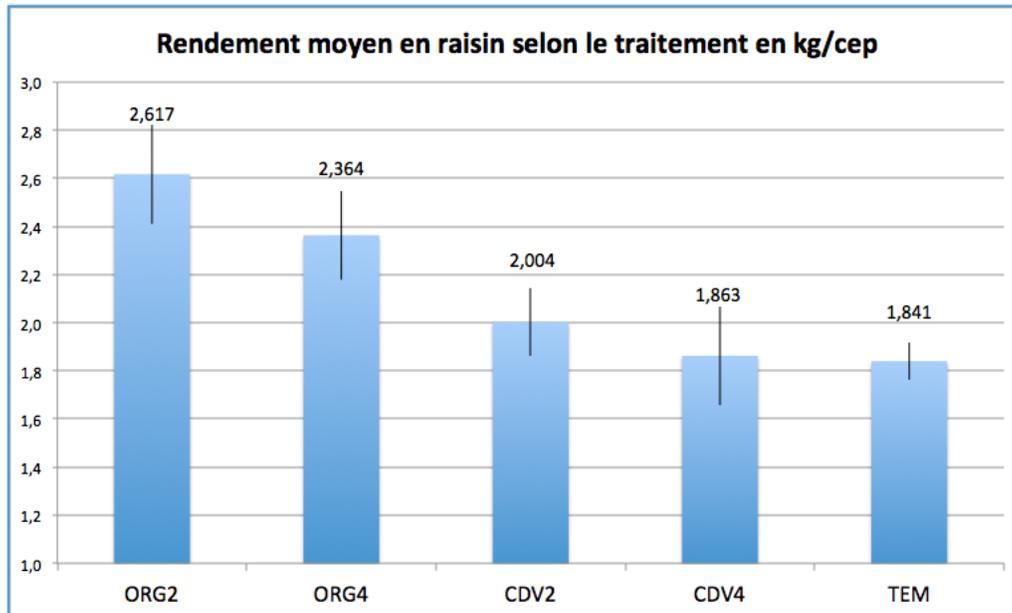


FIGURE 8 – Histogramme réalisé sur Statbox

L'histogramme ci-dessus montre des différences pour le paramètre considéré entre les traitements. Les traitements ORG2 et ORG4 se distinguent des autres traitements par un rendement moyen en raisin plus important. Les parcelles non traitées (traitement TEM) ainsi que les parcelles ayant reçu un apport de compost végétal (VEG2 et VEG4) ont des rendements moyens en raisin plus faibles, comparés aux rendements obtenus sur les parcelles ayant reçu un apport d'Orgaveg (ORG2 et ORG4). On peut raisonnablement supposer que les traitements ORG2 et ORG4 augmentent la productivité des ceps. Une ANOVA suivi d'un test de comparaison de moyennes permettra de valider ou non ce premier constat.

Pour réaliser des tests ANOVA, certaines hypothèses sur les résidus doivent être vérifiées.

3.1.1 Cartographie des résidus

Le logiciel Stabox effectue le calcul des résidus selon le modèle de l'Anova :
$$\text{Résidu} = \text{valeur mesurée} - (\text{effet moyen} + \text{effet bloc} + \text{effet traitement})$$

Le logiciel établit des classes de valeur des résidus et les cartographie sur le plan expérimental. La figure 9 représente la cartographie des résidus de l'essai St-Etienne pour le paramètre rendement moyen par cep en raisin (estimé le 13 septembre 2010), avec les résidus de chaque parcelle élémentaire.

Cartographie des résidus :

	TEM	ORG2	ORG4	VEG2	VEG4
Bloc 1					
Bloc 2					
Bloc 3					
Bloc 4					

Légende :

	Donnée manquante
	< - 0,106
	< 0,000
	< 0,106
	< 999999,000

FIGURE 9 – Cartographie des résidus de l’essai Saint-Etienne au 13 Septembre 2010

Au vu de la cartographie des résidus, l’expérimentateur vérifie un éventuel regroupement de parcelles dont les résidus présentent la même couleur (couleur allant du blanc au bleu foncé). Cela serait là le signe d’un potentiel biais expérimental.

Avec la cartographie des résidus, il est également possible de vérifier visuellement la présence ou non de résidus suspects, ou de gradient intra-bloc, ce qui pourrait introduire un biais expérimental. Plus la couleur est foncée, plus grand est le résidu. À l’inverse, les parcelles dont les résidus sont représentés en blanc présentent les valeurs de rendement les plus faibles. Ici, l’absence de regroupement d’une même couleur des résidus sur la cartographie nous incite à valider l’hypothèse de résidus répartis aléatoirement, et de l’absence de biais expérimental sur cette variable.

3.1.2 Égalité des variances dans les traitements et dans les blocs

L’analyse de variance des résidus exige d’avoir une égalité des variances du paramètre pour les différents traitements. L’hypothèse d’égalité des variances inter-bloc et intra-traitements est ainsi testée.

Ecart-types du facteur Traitement		Ecart-types du facteur Bloc	
Traitement	E.T.		E.T.
1 (1-TEM)	0,244	1 (B1)	0,221
2 (2-ORG2)	0,129	2 (B2)	0,134
3 (3-ORG4)	0,069	3 (B3)	0,098
4 (4-CDV2)	0,176	4 (B4)	0,195
5 (5-CDV4)	0,201		
$khi^2 = 3,953$ Prob. = 0,41332		$khi^2 = 2,677$ Prob. = 0,44628	

TABLE 3 – Test de Bartlett de l’homogénéité des variances dans les traitements et dans les blocs

Visuellement, il ne semble pas y avoir de différence flagrante entre les écart-types des différents traitements et blocs. Les statistiques de test associées au test d’égalité des traitements et des blocs valent respectivement 3,95 et 2,68. Les p-valeurs sont bien au dessus du seuil de significativité 0,05, on admet l’égalité des variances pour chacun des groupes définis par les traitements et par les blocs.

3.1.3 Normalité des résidus

La Figure 10 présente l’histogramme des résidus de chaque parcelle élémentaire. Il permet d’identifier les résidus pouvant être problématiques. De façon à pouvoir repérer facilement les observations correspondant aux résidus, chaque parcelle est identifiée sur l’histogramme par son numéro (voir figure 10). Par exemple, le résidu qui a le code 105 (point culminant de l’histogramme) est situé dans le bloc 1 et a reçu le traitement 5, c’est-à-dire le compost de déchets verts tous les 4 ans. Le nombre de classes (4 ici) est calculé par Statbox de telle sorte que l’utilisateur puisse avoir une image optimale de la distribution des résidus : pour n mesures, le nombre de classes sera proche de \sqrt{n} .

9		105		
8		304		
7		204		
6		403		
5	405	303	305	
4	104	103	205	
3	202	302	404	
2	401	102	203	
1	201	301	402	101
Effectifs				
	5	9	5	1
Bornes				
	-0,26	-0,11	0,05	0,21
	à	à	à	à
	-0,11	0,05	0,21	0,36

FIGURE 10 – Histogramme des résidus de l’essai Saint-Etienne

L’histogramme semble proche d’une distribution normale, on ne rejette pas l’hypothèse nulle de normalité des résidus.

Un test du Chi-deux d’ajustement à la loi normale aurait été possible si les résidus avaient été nombreux (pour un essai avec davantage de parcelles).

Le coefficient d’asymétrie de Pearson, indice de normalité des résidus, vaut 0,177. Cette valeur est suffisamment proche de zéro pour que l’on puisse considérer que la distribution est symétrique.

Le coefficient d’aplatissement de Pearson vaut 2,881. Cette valeur est suffisamment proche de 3 pour valider l’hypothèse que la distribution des résidus a le même aplatissement qu’une loi normale.

L’ensemble des conditions d’application de l’ANOVA pour le paramètre Y ont donc été vérifiées. Il est maintenant possible de procéder à l’analyse de variance.

3.2 Test global de l’effet des facteurs traitement et bloc

On veut savoir si le rendement moyen en raisin par cep varie différemment selon le traitement reçu, respectivement pour valider *a posteriori* le dispositif pour la variable considérée (effet bloc) et vérifier s’il y a une différence significative du facteur traitement sans renseigner sur les traitements significativement différents (effet traitement). Cette analyse est faite à l’aide du test F-Student.

L’ETR de l’essai est de 0,195 kg de raisin. Statistiquement, cet ETR correspond à un essai précis pour la variable mesurée. D’un point de vue agronomique, cette valeur n’est pas surprenante par rapport aux ETR des essais des années précédentes : on valide donc l’essai agronomiquement et statistiquement.

On vérifie dans un premier temps qu’il n’existe pas d’interaction facteur \times traitement. Le test d’inter-

action de Tukey (interaction bloc \times traitement) donne une p-valeur de 0,506, on se place donc dans le modèle sans interaction bloc \times traitement.

	S.C.E	DDL	C.M.	TEST F	PROBA
Var.TOTALE	2,633	19	0,139		
Var.FACTEUR 1	1,851	4	0,463	12,129	0,000
Var.BLOCS	0,324	3	0,108	2,828	0,083
VAR.RESIDUELLE	0,458	12	0,038		

TABLE 4 – Table Statbox d’analyse de la variance à deux facteurs

Les résultats du test F présentés dans le tableau 4 indiquent qu’il existe de façon significative une différence entre les traitements pour le paramètre Y (p-valeur $< 0,05$). En revanche, le test F-student montre que la probabilité critique de l’effet bloc est de 0,083 ce qui est supérieur à 0,05. L’effet bloc n’est pas significatif au niveau 5%, mais il le serait si l’expérimentateur était prêt à accepter un risque de première espèce de 10%. ce résultat indique que le terrain serait homogène, mais les blocs seraient mal répartis.

La puissance de l’essai est de 80%, ce que l’on considère comme satisfaisant : l’essai est bien assez puissant pour mettre en évidence des différences entre les traitements.

3.3 Comparaison entre les traitements

Il s’agit maintenant d’évaluer si les traitements présentent des différences significatives entre eux, en utilisant le test de Newman-Keuls.

On désigne par μ_{TEM} , μ_{ORG2} , μ_{ORG4} , μ_{VEG2} et μ_{VEG4} les moyennes du paramètre observé *Rendement moyen en raisin* pour les cinq traitements du dispositif (variable estimée le 13 Septembre 2010). On classe les moyennes des traitements dans l’ordre décroissant :

$$\mu_{\text{ORG2}} > \mu_{\text{ORG4}} > \mu_{\text{VEG4}} > \mu_{\text{VEG2}} > \mu_{\text{TEM}}.$$

Les moyennes sont présentées dans le tableau suivant.

Traitement	μ_{ORG2}	μ_{ORG4}	μ_{VEG4}	μ_{VEG2}	μ_{TEM}
Moyenne du rendement moyen en raisin	2,617	2,364	2,004	1,863	1,841

TABLE 5 – Moyennes du rendement moyen en raisin obtenues pour chacun des traitement de l’essai St-Etienne

L’hypothèse nulle est (H0) : $\mu_i = \mu_j$.

Cette hypothèse est rejetée si la différence entre ces 2 moyennes est plus grande que la plus petite amplitude significative (PPAS). L’expression de la PPAS est donnée ci-dessous :

$$PPAS = t_s \sqrt{\frac{CMR}{q}}$$

CMR correspond au carré moyen résiduel, sa valeur se lit dans n’importe quelle table d’ANOVA.

q est le nombre de répétitions d’un traitement.

t_s est fonction du nombre k de moyennes à regrouper, du degré de liberté de la variance résiduelle et d’un niveau de risque δ que l’on choisit. On lit sa valeur dans une table Range Studentisé.

Nous allons successivement comparer 2, puis 3, puis 4 et enfin 5 moyennes entre-elles, au risque $\delta = 0,05$ de se tromper en rejetant l'hypothèse nulle. La table d'ANOVA précédente donne $CMR = 0,038$ (à la dernière ligne, colonne 4) et 12 degrés de liberté pour la variance résiduelle. Chaque traitement est répété 4 fois. Ces informations et la table Range Studentisé disponible en annexe nous permettent d'obtenir le tableau suivant.

Nombre de moyennes	2	3	4	5
t_s	3,08	3,77	4,2	4,51
PPAS	0,3	0,37	0,41	0,44

TABLE 6 – Plus petite amplitude significative pour 2, 3, 4 et 5 moyennes comparées

On déduit de ce tableau un deuxième tableau 7, qui va nous permettre d'effectuer manuellement le test de Newman-Keuls.

Traitement	μ_{ORG2}	μ_{ORG4}	μ_{VEG4}	μ_{VEG2}	μ_{TEM}
Moyennes	2,617	2,364	2,004	1,863	1,841
Différences ORG2 - autres traitements	0	0,253	0,613	0,754	0,776
PPAS		0,300	0,367	0,409	0,440
Groupe 1	OUI	OUI	NON	NON	NON
Différences ORG4 - autres traitements		0	0,360	0,501	0,523
PPAS			0,300	0,367	0,409
Groupe 2		OUI	NON	NON	NON
Différences VEG4 - autres traitements			0	0,141	0,022
PPAS				0,300	0,367
Partie du Groupe 2			OUI	OUI	OUI

TABLE 7 – Test de Newman-Keuls de distinction de groupes homogènes parmi 5 traitements

Les deux premières lignes rappellent les différents traitements et leur rendement moyen en raisin par cep. La première comparaison (troisième ligne) compare la moyenne la plus élevée, à savoir μ_{ORG2} avec toutes les autres moyennes. Les valeurs de cette ligne correspondent à la différence entre μ_{ORG2} et les autres moyennes. À titre d'exemple, le 0,253 s'obtient par la différence entre $\mu_{ORG2} = 2,617$ et $\mu_{ORG4} = 2,364$

La ligne suivante présente les plus petites amplitudes significatives associées aux comparaisons respectives de 2, 3, 4 et 5 moyennes.

À la ligne 5, on compare les différences entre 2 moyennes à la PPAS; tant que la différence entre deux moyennes reste inférieure à la PPAS, l'hypothèse nulle n'est pas rejetée et les deux traitements comparés font partie du même groupe homogène. Par exemple, on sait dès cette première ligne que le traitement ORG2 (apport d'Orgaveg tous les 2 ans) et le traitement ORG4 (apport d'Orgaveg tous les 4 ans) feront partie d'un même groupe de traitement à la fin du test.

On réitère le processus jusqu'à obtenir au moins un "OUI" dans chaque colonne.

Deux groupes homogènes se distinguent : un groupe A constitué des traitements ORG2 et ORG4, et un second groupe B constitué des traitements CDV4, CDV2 et TEM (tableau 8). Ce résultat confirme la tendance observée dans l'histogramme de la figure 8.

Le test de Newman-Keuls effectué par Statbox confirme ce résultat avec les résultats présentés dans le tableau 8 suivant, les rendements moyens en raisin des traitements ORG2 et ORG4 sont

significativement supérieurs aux rendement moyens en raisin des traitement VEG2, VEG4 et TEM.

Groupes homogènes			
Id	Modalité	Moyenne	Groupes homogènes
2	T2 ORG2	2,617	A
3	T3 ORG4	2,364	A
5	T5 CDV4	2,004	B
4	T4 CDV2	1,863	B
1	T1 TEM	1,841	B

TABLE 8 – Test de Newman-Keuls de comparaison de traitements

D'un point de vue agronomique, on peut ainsi conclure au fait que l'apport du PRO Orgaveg tous les 2 ans ou tous les 4 ans permet d'augmenter significativement le rendement moyen en raisin par cep par rapport à un apport de compost de déchets tous les 2 ou 4 ans, et par rapport au traitement témoin.

3.4 Problèmes rencontrés et solutions statistiques

Nous avons vu les manières de procéder lorsque toutes les conditions nécessaires pour faire l'ANOVA et les tests sont réunies. Mais en pratique, il arrive parfois que certains problèmes soient rencontrés au cours de la procédure statistique. Il existe des solutions statistiques permettant de s'affranchir de ces difficultés. Les obstacles les plus souvent rencontrés et les réponses statistiques sont présentés dans les paragraphes suivants.

3.4.1 Présence de résidus suspects

Comme nous l'avons vu, il est possible de repérer une valeur statistiquement aberrante avec le test de Grubbs. Cependant, l'identification d'une valeur aberrante par le test de Grubbs doit être validée agronomiquement par l'expertise de l'agronome. Le test de Grubbs est ainsi appliqué dans le but d'attirer l'attention sur une valeur jugée statistiquement aberrante. Dans le cas où la valeur statistiquement aberrante est également considérée comme aberrante d'un point de vue agronomique, il est possible de la remplacer par une valeur estimée par le modèle linéaire de Yates.

Nous avons vu que toute mesure peut s'écrire sous la forme $Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$, soit l'addition d'un effet global, d'un effet lié au traitement, d'un effet du bloc et un effet résiduel. Une valeur aberrante (outlier) observée Y_{ij} s'éloigne de la valeur théorique attendue lorsqu'elle est caractérisée par un résidu ϵ_{ij} anormalement élevé ou faible. S'il n'y a aucune justification agronomique valable pour l'expliquer, il faut la remplacer par sa valeur théorique \widehat{Y}_{ij} . C'est le principe de l'estimation de Yates.

La valeur estimée est $\widehat{Y}_{ij} = \widehat{\mu} + \widehat{\alpha}_i + \widehat{\beta}_j$.

Ces paramètres sont respectivement estimés par les grandeurs suivantes :

$\widehat{\mu}$ est la moyenne empirique, calculée pour le paramètre sur toutes les parcelles expérimentales ;

$\widehat{\alpha}_i$ se calcule comme la différence entre la moyenne empirique du paramètre sur le traitement i et $\widehat{\mu}$;

$\widehat{\beta}_j$ s'obtient comme la différence entre la moyenne empirique des rendements du bloc j et $\widehat{\mu}$;

Dans le cas où il y a plus de 10% de valeurs manquantes/aberrantes, l'expérimentateur n'a d'autre choix que d'invalider l'essai sur le paramètre et pour la date de prélèvement considérée (M. Vilain).

3.4.2 Non normalité

La normalité des résidus peut être remise en cause pour les raisons suivantes :

- l'histogramme met en évidence une répartition non normale des résidus, avec plusieurs classes modales éloignées par exemple, ou met en évidence la présence de résidu(s) suspect(s) ;
- le coefficient d'asymétrie de Pearson est éloigné de la valeur 0 attendue ;
- le coefficient d'aplatissement de Pearson est éloigné de la valeur 3 attendue ;
- présence de résidu(s) suspect(s).

Si l'on est dans une de ces situations, il peut être utile d'opérer une transformation sur la variable afin de stabiliser les variances.

Selon M. Vilain (1999), il faut substituer à la variable étudiée une variable de travail, une variable transformée, choisie de manière à mieux satisfaire aux conditions d'application.

On récapitule les transformations possibles et leurs motivations dans le tableau suivant.

Transformation	Justifications
$Y = \sqrt{X}$	La variable X est discontinue Sa distribution est voisine de celle de Poisson $\mu_X = \sigma_X^2$
$Y = \sqrt{X + 1}$ ou $Y = \sqrt{X}\sqrt{X + 1}$	Idem que pour la transformation racine carrée, mais avec des valeurs de X faibles
$Y = \ln X$	Les moyennes par bloc ou par traitement sont proportionnelles aux écart-types dans les bloc ou dans les traitements
$Y = \ln(X + 1)$	Les valeurs mesurées sont faibles Les valeurs mesurées sont des entiers
$Y = \arcsin X$	X semble suivre une loi binomiale X est exprimée en pourcentage

TABLE 9 – Les différentes transformations de variable possibles et les justifications correspondantes

Intéressons nous au paramètre *Poids moyen du bois de taille* (BDT), prélevé pour l'essai St-Etienne le 9 mars 2009 et exprimé en kg/cep. Les résidus ne sont visiblement pas distribués normalement, comme en témoigne l'histogramme ci-dessous.

Histogramme des résidus :

8	405			
7	205			
6	303			
5	203	404		
4	202	304	305	
3	102	204	105	103
2	401	104	403	402
1	101	301	302	201
Effectifs				
	8	5	4	3
Bornes				
	-0,04	-0,01	0,01	0,04
	à	à	à	à
	-0,01	0,01	0,04	0,07

FIGURE 11 – Histogramme des résidus du poids moyen BDT non normal au 9 mars 2009

La distribution du poids moyen BDT est sensiblement proche d'une distribution de Poisson. La démarche à adopter est donc une transformation racine carré de la variable. Toutefois, après transformation racine carrée, la distribution n'est toujours pas normale.

Lorsque les transformations de variables vues précédemment n'ont pas donné de résultat, il est encore possible de faire un test non paramétrique. Un test non paramétrique basé sur les rangs des observations ne fait aucune hypothèse sur la normalité de la variable observée, il est donc possible de les appliquer. Toutefois, les tests non paramétriques sont jugés moins puissant que les tests paramétrique comme l'ANOVA, et ne doivent être utilisés qu'en dernier recours, lorsque les tests paramétriques ne peuvent être appliqués.

Le test de Kruskal et Wallis (1952) permet de savoir si k échantillons indépendants sont issus de la même population. C'est une généralisation du test de Wilcoxon-Mann-Whitney à plus de 2 échantillons.

Considérons le paramètre phosphore total (P), mesuré sur la couche de sol 0-15 cm au 10 février 2009 pour l'essai St-Etienne. Sa distribution présentait un défaut de normalité que les transformations de variables n'ont pas pu corriger. Un test non paramétrique de Kruskal-Wallis se substitue alors à l'ANOVA. Les données sont présentées dans le tableau ci-dessous.

TEM	ORG2	ORG4	VEG2	VEG4
1,1	1,18	1,27	1,11	1,32
0,97	0,91	1,19	1,07	1,08
1,04	0,92	1,09	1,01	1,07
0,98	0,93	0,94	0,89	1,12

TABLE 10 – Présentation des mesures du paramètre P total HF par traitement au 10 février 2009

On dispose de $p = 5$ échantillons de taille $n_i = 4$. Un échantillon correspond à un traitement épandu. Il s'agit de comparer les valeurs du P total dans les 5 traitements, pour mettre en évidence un éventuel effet du traitement sur la teneur observée en P total dans la couche de sol 0-15 cm. On teste ainsi au risque $\delta = 0,05$

(H0) : *les 5 échantillons suivent la même distribution* contre

(H1) : *au moins un échantillon suit une distribution de position différente.*

On note r_{ij} le rang de la mesure prélevée sur la parcelle du bloc j qui a reçu le traitement i. Le rang 1 est affecté à la plus petite mesure, le rang n (20 ici) à la plus grande. Si des ex-aequo existent, on leur affecte des rangs moyens. On calcule alors les quantités suivantes :

$SDR_i = \sum_{j=1}^{n_i} r_{ij}$, la somme des rangs pour chaque traitement i ;

$SC = \sum_{i=1}^p \frac{SDR_i^2}{n_i}$, la somme des carrés générale ;

$SCR = \sum_{i=1}^p \sum_{j=1}^q r_{ij}^2$, la somme des carrés des rangs ;

$C = \frac{1}{4}n(n+1)^2$, une correction appropriée.

La statistique de test est définie par :

$$H = (n-1) \frac{SC - C}{SCR - C}$$

Sous l'hypothèse nulle, H suit approximativement un χ^2 à $p-1$ degrés de libertés. On compare cette valeur à une valeur critique $\chi_{p-1;\delta}$, lue dans la table du Chi-deux (disponible en annexe). On

rejette (H0) quand $H > \chi_{p-1;\delta}$, où $P(H > \chi_{p-1;\delta}) = \delta$.

Sous Excel, à l'aide des fonctions RANG et SOMME, on obtient le tableau suivant.

Traitement	P tot HF	RANG	RANG ²	SDR _i	SDR _i ² / n _i
T1	1,1	14	196		
T1	0,97	6	36		
T1	1,04	9	81	36	324
T1	0,98	7	49		
T2	1,18	17	289		
T2	0,91	2	4		
T2	0,92	3	9	26	169
T2	0,93	4	16		
T3	1,27	19	361		
T3	1,19	18	324		
T3	1,09	13	169	55	756,25
T3	0,94	5	25		
T4	1,11	15	225		
T4	1,07	10,5	110,25		
T4	1,01	8	64	34,5	297,5625
T4	0,89	1	1		
T5	1,32	20	400		
T5	1,08	12	144		
T5	1,07	10,5	110,25	58,5	855,5625
T5	1,12	16	256		

TABLE 11 – Tableau des étapes de calcul de la statistique H du test de Kruskal-Wallis

La valeur 1,07 apparaît à deux reprises, une fois dans le traitement 4, et une fois dans le traitement 5. On affecte donc à ces 2 mesures le rang moyen 10,5.

On déduit du tableau 11 la valeur des quantités précédentes.

La somme des rangs pour chaque traitement apparaît dans le tableau, avec :

SC, qui s'obtient en sommant les valeurs de la dernière colonne : $SC = 2402,375$;

SCR, qui s'obtient en sommant les valeurs de la colonne $RANG^2$: $SCR = 2869,5$;

$C = 2205$;

De ces quantités découle la valeur de la statistique de test : $H = 5,6435$ à 10^{-4} près. Sous (H0),

$H \sim \chi^2(5 - 1)$. Dans la table du Chi-deux en annexe, on trouve 9,488 comme valeur de $\chi_{4;0,05}$.

$H < \chi_{4;0,05}$, on ne rejette pas l'hypothèse que les traitements suivent la même distribution au risque 0,05.

Statbox calcule directement la p-valeur, mais fournit également la valeur de H, ainsi que celle de $\chi_{p-1;\delta}$. Cela nous permet de vérifier les résultats trouvés précédemment.

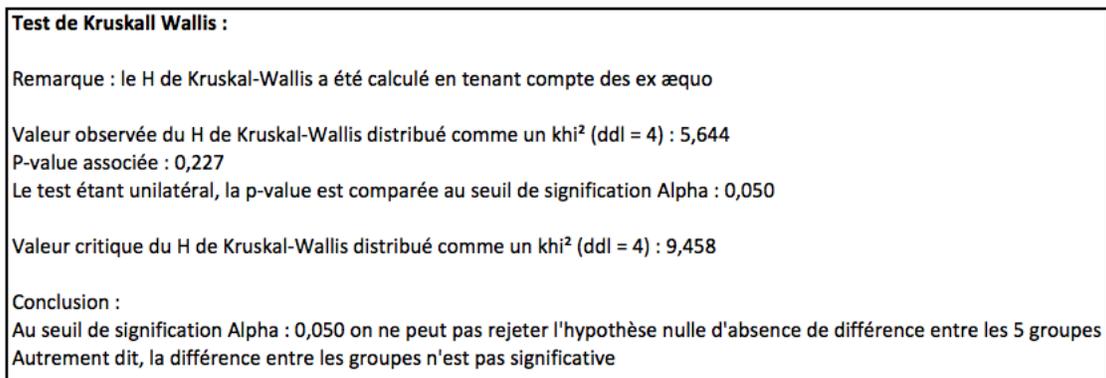


FIGURE 12 – Sortie Statbox du test de Kruskal-Wallis

La présence de valeurs égales est confirmée à la première ligne. La loi de H est bien un Chi-deux à 4 degrés de liberté. Si on arrondi à 10^{-3} la valeur de H trouvée, on a effectivement $H = 5,644$. Au seuil de signification 0,05, la valeur de $\chi_{4;0,05}$ trouvée par Statbox est 9,458. Elle diffère donc sensiblement de celle que l'on peut lire dans la table du Chi-deux. La conclusion du test demeure de toute façon inchangée, comme le confirme la p-valeur de 0,227 : cette valeur est supérieure au seuil 0,05, on ne dispose pas de preuve suffisante pour rejeter l'hypothèse (H_0). On admet que le paramètre P total est distribué identiquement dans les groupes de traitement. Il n'existe donc pas de différences significatives entre les traitements

J'ai procédé de la même façon à chaque fois qu'un défaut de normalité était détecté par Statbox pour les différentes variables.

3.4.3 Hétéroscedasticité

Parfois, les variances des résidus ne sont pas les mêmes d'un traitement ou d'un bloc à l'autre. Les expériences en blocs aléatoires complets possèdent l'avantage de conserver toutes leurs propriétés quand, *pour quelque raison que ce soit*, un ou plusieurs blocs doivent être éliminés de l'analyse des résultats (P. Dagnelie, 1981). Il est permis de supprimer le bloc ou le traitement de l'analyse, mais seulement s'il existe une justification agronomique à cette hétérogénéité. Une bonne raison pour retirer un bloc de l'analyse peut être une erreur d'apport, un dégât accidentel, un passage d'animaux, la présence d'un gradient intra-bloc observé sur la cartographie des résidus et confirmé par expertise, une interaction bloc \times traitement...

Conclusion

Au cours de mon stage, les données de 3 essais conduits dans le cadre du projet Réseau MO piloté par l'IFV ont été saisies dans un fichier Excel de saisie en vue d'être insérées dans la base de données du Réseau PRO. Pour chaque essai saisi, la moyenne et l'écart-type ont été calculés et saisis dans le fichier pour chaque variable de chaque traitement à chaque date de prélèvement. Ceci permettra de mettre de façon immédiate aux utilisateurs de la base de données du Réseau PRO des informations sur la qualité statistique des données disponibles et faciliter l'ouverture ultérieure des droits aux utilisateurs des données. Cependant la moyenne et l'écart-type sont des indicateurs succincts qui doivent être complétés par d'autres indicateurs de statistiques descriptives, et par des analyses statistiques permettant notamment de valider les jeux de données.

Les tests statistiques effectués sur les jeux de données des essais conduits en viticulture du Réseau MO ont permis d'avoir une meilleure connaissance de la validité statistique et agronomique des essais, pour chacun des paramètres mesurés. Outre l'estimation de la puissance de l'essai qui permet de dire si l'essai est acceptable pour mettre en évidence des différences significatives entre traitements, l'analyse de variance (si les conditions sont remplies) permet de mettre en évidence des différences significatives entre les différents traitements comparés, pour un paramètre considéré. Si des différences significatives ont été soulignées par ce test, il s'agit ensuite de les caractériser. Des tests de comparaison de moyennes, tel que le test de Newman-Keuls, permettent de distinguer des groupes homogènes de traitements significativement différents les uns des autres. Ces tests statistiques permettent de tirer des conclusions agronomiques sur les phénomènes étudiés et les facteurs les influençant. Ils peuvent en outre permettre d'identifier des problèmes expérimentaux, tels qu'un manque de précision de l'essai (ETR trop élevé) qui pourraient être palliés en augmentant par exemple le nombre de répétitions des traitements.

Dans le cas de l'essai de Saint-Etienne la Varenne présenté ici, l'ensemble de ces tests a permis, par exemple, de montrer qu'il existe une différence significative des traitements testés sur le rendement moyen par cep en raisin. D'un point de vue agronomique, on a donc pu conclure que, dans les conditions de l'expérimentation, l'apport du PRO « Orgaveg » tous les 2 ans ou tous les 4 ans permet d'obtenir un rendement moyen par cep significativement plus important que lorsqu'aucun apport organique n'est fait (traitement témoin) ou que lorsque l'on apporte du compost végétal.

Dans certains cas, des problèmes d'application des tests statistiques peuvent être rencontrés. Selon la nature de ces problèmes, plusieurs solutions statistiques permettent de les contourner.

Certaines valeurs peuvent être aberrantes. Cela se traduit par la présence de résidus suspects. Il est alors possible de corriger la valeur par la méthode de Yates. Ou encore, comme cela a été le cas pour le poids des bois de taille, une distribution non normale d'un paramètre s'oppose à l'analyse de variance. Une transformation de la variable en question permet parfois d'aboutir à une distribution normale. On peut alors effectuer l'ANOVA et les tests de comparaison de moyenne sur la variable transformée. Dans le cas où une transformation ne donne pas de résultat, on peut alors effectuer un test non paramétrique tel que le test de Kruskal-Wallis. Ce test permet de mettre en évidence ou non l'existence de différence significative d'une variable entre les traitements comparés. Ce test a été effectué sur le poids des bois de taille dont la distribution n'était pas normale, même après

transformation. Il n'a cependant pas permis de mettre en évidence de différences entre les traitements pour ce paramètre. D'autres problèmes peuvent être rencontrés tels que des interactions entre les effets blocs et les effets des traitements, des gradients existants au sein d'un bloc pourtant supposé homogène ou des inégalités de variance des résidus. L'élimination d'un résidu suspect ou d'un bloc permet dans certains cas de s'affranchir de ces problèmes et de poursuivre l'analyse statistique des données. Dans le cas où aucune solution statistique n'apporte de résultats pour un paramètre donné, il est donc impossible de poursuivre l'analyse statistique et de conclure sur l'existence d'effets ou non des traitements comparés dans l'essai. D'autre part, comme les résultats des tests effectués sur le rendement moyen par cep et sur le poids des bois de taille le montrent, il est possible qu'un essai agronomique soit puissant pour un paramètre mesuré donné et ne le soit pas pour un autre. C'est pourquoi, au-delà d'une interprétation statistique des données, il est important d'apporter une explication agronomique aux résultats statistiques des tests effectués.

Les données que j'ai saisies seront par la suite insérées dans la base de données du Réseau PRO. Tester les jeux de données était donc l'assurance d'avoir une transparence sur leurs qualités agronomiques et statistiques pour des paramètres donnés. Cette transparence est importante car les données saisies et insérées dans la base de données du Réseau PRO seront mises à disposition à l'ensemble des partenaires du Réseau PRO. Les tests statistiques effectués, couplés à une interprétation agronomique, permettront aux experts du Réseau PRO de mettre au point un indicateur de qualité pour chaque paramètre mesuré sur un essai.

Sur un plan plus personnel, ce stage m'a permis de découvrir le milieu agronomique, que je ne connaissais pas. J'ai vu comment et pourquoi la statistique pouvait s'intégrer dans ce milieu. J'ai compris la démarche des expérimentateurs quant à la mise en place des essais agronomiques, à la formulation d'hypothèses, au choix du dispositif expérimental et à l'interprétation des résultats. Grâce à la visite d'un essai agronomique en plein champ (« QualiAgro » Feucherolles (78)), je me suis familiarisé avec les notions d'essai et de dispositif en blocs.

Ce stage a complété ma formation à l'Institut Universitaire de Technologie. J'ai acquis des connaissances nouvelles, dans le domaine des tests d'hypothèses notamment (Newman-Keuls, Kruskal-Wallis). J'ai appris à me servir d'un logiciel non rencontré au cours de ma formation (Statbox). Et surtout, j'ai pu mettre en pratique, sur des situations concrètes, les enseignements d'analyse de variance et de tests paramétriques et non paramétriques dispensés en D.U.T.

Bibliographie

Dans leur ordre d'apparition, voici les différentes sources dont je me suis inspiré pour la rédaction du présent rapport.

Michel Vilain, *Méthodes expérimentales en agronomie – Pratiques et analyses*, 1999

Philippe Letourmy, *Expérimentation agronomique planifiée*, 1999

Jean Pierre Gouet et Gérard Philippeau, *Comment interpréter les résultats d'une analyse de variance ?*, 2002

Maurice Stevenson Bartlett, *Properties of sufficiency and statistical tests*, 1937

Carletti Grubbs, *Détection automatique de valeurs aberrantes*, 1975

Peter Sprent, *Pratique des statistiques non paramétriques*, 1992

Pierre Dagnelie, *Principes d'expérimentation – Planification des expériences et analyse de leurs résultats*, 1981

Ressources internet :

- Le site de l'Institut Français de la Vigne et du Vin
- L'ancien site de l'Université Claude Bernard à Lyon
- Le site du logiciel Statbox

Fiches synoptiques des essais Saint-Etienne, Orgueil et Vauvert

Système de production : Conventionnel	N° Fiche :
Effets de divers apports de matière organique sur le sol, la vigne et le vin : Saint Etienne La Varenne	Viticulture
	Répétitions classiques (4)

INFORMATIONS GENERALES	Maître d'ouvrage : InterBeauJolais Maître d'œuvre : SICAREX Beaujolais	Programme rattaché à l'essai : Réseau national MO IFV	Département : Rhône Commune : Saint Etienne La Varenne
	Service à contacter : IFV Villefranche-sur-Saône (Cahurel Jean-Yves) 210 en Beaujolais BP 320 69661 Villefranche-sur-Saône cedex Tél.: 04 74 06 43 43 E-mail : Jean-Yves.CAHUREL@vignevin.com	Partenaires techniques IFV Partenaires scientifiques ENESAD, AgroTransfert Partenaires financiers	

Contexte et objectifs de l'essai	
Amélioration du conseil en termes de matière organique auprès des viticulteurs Paramétrage et validation du modèle AMG sur vigne Connaissance des produits testés	

CONTEXTE EXPERIMENTAL	Début de l'essai 2009 Fin de l'essai 2016 Durée 8 ans	Contexte pédologique - Type pédologique de sol : - Substrat pédologique : Arène - Nature du sous-sol : - Texture du sol : A = 7,9%, L = 17,2%, S = 74,9% - Profondeur de sol : non connue, > 170 cm - pH initial horizon de surface : 5,8 - Autres caractéristiques : Teneur en MO initiale de 0,8%	Type de dispositif : Blocs Surface totale du dispositif : 3326 m ² Surface des parcelles élémentaires : 64,7 m ² Facteurs étudiés : Nature du PRO (2), Fréquence d'apport du PRO (2) Type de témoin : inclus au dispositif, fixe, sans apport fertilisant Nombre de traitements : 5 Nombre de répétitions par traitement : 4
	Type de PRO testés		Système de culture
	Compost de déchets verts Amendement organique du commerce		Cépage : Gamay Plantation des ceps : 2002

SUIV DES COMPARTIMENTS	PRO	Physico-chimiques : MS, C _{org} , MO, N _{Total} , P, K, Ca, Mg Fractionnement biochimique (ISMO)		
	Sol	<i>Inter-rang</i>	Physico-chimiques : pH, C _{org} , MO, N _{Total} , P, K, Ca, Mg, CEC, Cations échangeables Granulométrie 5 fractions, ETM : Cu, Mn	
	Plante	<i>Rognages, Bois de taille, Feuilles</i>	Production de biomasse par cep Teneurs en C et N	
		<i>Pétiole</i>	Teneurs en K, Mg	
		<i>Fruits</i>	Rendement : poids de la récolte par cep, poids moyen de fruit/grappe, poids des baies	
	<i>Moût</i>	Qualité : teneur en sucre, pH, acidité, azote assimilable		

<i>Sources des informations fournies :</i> IFV Villefranche-sur-Saône <i>Type d'organisme :</i> Institut technique	<i>Format des données de l'essai :</i> Fichiers Excel – BDD IFV <i>Mode de diffusion des résultats :</i> <i>Niveau de confidentialité :</i>
---	---

Système de production : Conventionnel	N° Fiche :
Effets de divers apports de matière organique sur le sol, la vigne et le vin : Orgueil	Viticulture
	Répétitions classiques (3)

INFORMATIONS GENERALES	Maître d'ouvrage : IFV Maître d'œuvre : IFV Sud-Ouest	Programme rattaché à l'essai Réseau National MO IFV	Département : Tarn-et-Garonne Commune : Orgueil
	Service à contacter : IFV Sud-Ouest (Laure Gontier) V' innopôle, BP22 81310 Lisle sur Tarn Tél. : 05 63 33 62 62 E-mail : laure.gontier@vignevin.com	Partenaires techniques IFV Partenaires scientifiques ENESAD, AgroTransfert Partenaires financiers	

Contexte et objectifs de l'essai	
Amélioration du conseil en termes de matière organique auprès des viticulteurs Paramétrage et validation du modèle AMG sur vigne Connaissance des produits testés	

CONTEXTE EXPERIMENTAL	Début de l'essai 2008 Fin de l'essai 2016 Durée 8 ans	Contexte pédologique - Type pédologique de sol : Luvisol redoxisol - Substrat pédologique : Alluvion - Nature du sous-sol : - Texture du sol : A : 12,3% ; L : 45,3% ; S : 40,6% - pH initial horizon de surface : 7,2 - Autres caractéristiques : indice de battance initial de 1,69, Teneur initiale en C de 1,04%	Type de dispositif : Blocs Surface du dispositif : 1080 m ² Surface des parcelles élémentaires : 72 m ² Facteurs étudiés : Nature du PRO (3), Enherbement (2) Type de témoin : inclus au dispositif, fixe, sans apport fertilisant Nombre de traitements : 5 Nombre de répétitions par traitement : 3
	Type de PRO testés		Système de culture
	Compost déchets verts Compost de marc de raisins Amendement organique du commerce		Cépage : Négrette Plantation des ceps : 1970

SUIVI DES COMPARTIMENTS	PRO	Physico-chimique : MS, C _{org} , MO, N _{Total} , P, K, Ca, Mg Fractionnement biochimique , Minéralisation C (3 jours)		
	Sol	Inter-rang	Physico-chimique : pH, CaCO ₃ , C _{org} , MO, N _{Total} , P, K, Ca, Mg, CEC Granulométrie 5 fractions, Taux d'éléments grossiers (refus à 2 mm) ETM totaux et extractibles : Mn, Cu Biomasse microbienne Densité apparente	
	Plante	Rognage, Bois de taille	Rendement (poids par cep) Physico-chimique : MS, C, N	
		Fruit	Rendement (poids par cep)	
		Pétiole	Physico-chimique : K, Ca, Mg	
	Moût	Qualité : Teneur en sucre, pH, Acidité, N assimilable		

Sources des informations fournies : IFV Sud-Ouest Type d'organisme : Institut technique	Format des données de l'essai : Fichiers Excel, BDD IFV Mode de diffusion des résultats : Niveau de confidentialité :
--	---

Système de production : Biologique		N° Fiche :
Effets de divers apports de matière organique sur le sol, la vigne et le vin : Vauvert		Viticulture
		Répétitions classiques (3)

INFORMATIONS GENERALES	Maître d'ouvrage : IFV Maître d'œuvre : IFV Nîmes	Programme rattaché à l'essai : Réseau National MO IFV	Département : Gard Commune : Vauvert
	Service à contacter : IFV Nîmes (Delpuech Xavier) Domaine de Donadille 30230 Rodilhan Tél. : 04 66 20 67 00 E-mail : xavier.delpuech@vignevin.com	Partenaires techniques IFV Partenaires scientifiques ENESAD, AgroTransfert	

Contexte et objectifs de l'essai	
Amélioration du conseil en termes de matière organique auprès des viticulteurs Paramétrage et validation du modèle AMG sur vigne Connaissance des produits testés	

CONTEXTE EXPERIMENTAL	Début de l'essai 2008 Fin de l'essai 2016 Durée 8 ans	Contexte pédologique - Type pédologique de sol : Fersialsol Eluvique (réf. Pédo. Français) - Substrat pédologique : - Nature du sous-sol : - Texture du sol : Limono-sableuse - Profondeur de sol : - pH initial horizon de surface : 7,35 - Autres caractéristiques : teneur en MO initiale de 1,2%	Type de dispositif : Carré latin Surface du dispositif : 1620 m ² Surface des parcelles élémentaires : 180 m ² Facteurs étudiés : Nature du PRO (2) Type de témoins : inclus au dispositif, fixe, sans apport fertilisant Nombre de traitements : 3 Nombre de répétitions par traitement : 3
	Type de PRO testés		Système de culture
	Compost déchets verts Compost de marc de raisin		Cépage : Syrah Plantation des ceps : 1998

SUIVI DES COMPARTIMENTS	PRO	Physico-chimiques : MS, C _{org} , MO, N _{Total} , P, K, Ca, Mg Fractionnement biochimique (ISMO) Minéralisation C (3 jours)		
	Sol	<i>Inter-rang</i>	Physico-chimiques : pH, C _{org} , MO, N _{Total} , P, K, Ca, Mg, CEC, Cations échangeables Granulométrie 5 fractions ETM : Cu, Mn Densité apparente	
	Plante	<i>Rognages, Bois de taille, Feuilles</i>	Production de biomasse par cep Teneurs en C et N	
		<i>Pétiole</i>	Teneur en K, Ca, Mg	
		<i>Fruit</i>	Rendement : poids de la récolte par cep, poids moyen de fruit/grappe, poids des baies	
	<i>Moût</i>	Qualité : Teneur en sucre, pH, Acidité, N assimilable		

Sources des informations fournies : IFV Nîmes Type d'organisme : Institut technique	Format des données de l'essai : Fichiers Excel, BDD IFV Mode de diffusion des résultats : Fiches d'information Niveau de confidentialité :
--	--

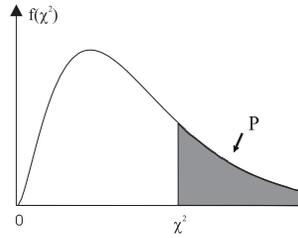
Tables de la distribution du Range Studentisé et du Chi-deux

Range Studentisé									
<p>k = nombre de moyennes, n = nombre de degrés de liberté de la variance résiduelle, $\alpha = 5\%$. La table donne la valeur de t_s seul intervenant dans le calcul des Plus Petites Amplitudes Significatives (test de Newmann-Keuls).</p>									
$n \backslash k$	2	3	4	5	6	7	8	9	10
1	18,0	27,0	32,8	37,1	40,4	43,1	45,4	47,4	49,1
2	6,08	8,33	9,80	10,9	11,7	12,4	13,0	13,5	14,0
3	4,50	5,91	6,82	7,50	8,04	8,48	8,85	9,18	9,46
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99
6	3,46	4,34	4,90	5,30	5,63	5,90	6,12	6,32	6,49
7	3,34	4,16	4,68	5,06	5,36	5,61	5,82	6,00	6,16
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92
9	3,20	3,95	4,41	4,76	5,02	5,24	5,43	5,59	5,74
10	3,15	3,88	4,33	4,65	4,91	5,12	5,30	5,46	5,60
11	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,39
13	3,06	3,73	4,15	4,45	4,69	4,88	5,05	5,19	5,32
14	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25
15	3,01	3,67	4,08	4,37	4,59	4,78	4,94	5,08	5,20
16	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15
17	2,98	3,63	4,02	4,30	4,52	4,70	4,86	4,99	5,11
18	2,97	3,61	4,00	4,28	4,49	4,67	4,82	4,96	5,07
19	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04
20	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01
24	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92
30	2,89	3,49	3,85	4,10	4,30	4,46	4,60	4,72	4,82
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,63	4,70
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65
120	2,80	3,36	3,68	3,92	4,10	4,24	4,36	4,47	4,56
∞	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47

TABLE 12 – Table de Newman-Keuls donnant la valeur de t_s

Table de la loi du Khi-deux

Valeurs de χ^2 ayant la probabilité **P** d'être dépassées



ν	P = 0,995	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01	0,005
1	0,00004	0,0002	0,001	0,0039	0,0158	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,39	10,865	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672

Nota. ν est le nombre de degrés de liberté.

Pour $\nu > 30$, on peut admettre que la quantité $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ suit la loi normale centrée réduite.