



Effets de l'azote et du système de culture sur la Pyriculariose du riz : analyse à l'aide du modèle linéaire généralisé

Stagiaire :

Emilie DOGE

Master 1 – IUP SID

2012/2013

Stage effectué au :

CIRAD de Montpellier
Avenue Agropolis
TA 178/04,
34398 Montpellier Cedex 5

Maître de stage :

M. Eric GOZÉ

Tuteur de stage :

Mme Cécile CHOUQUET

RÉSUMÉ

Du 2 Avril au 31 Juillet 2013, j'ai effectué mon stage au CIRAD (Centre de coopération internationale en recherche agronomique pour le développement) de Montpellier, au sein du département Persyst (Performances des systèmes de production et de transformation tropicaux), dans l'unité de recherche SCA (Systèmes de culture annuels). Durant ce stage j'ai travaillé sur une des principales maladies du riz pluvial à Madagascar : la pyriculariose.

En effet à Madagascar, le riz est l'aliment de base de la population. Le riz pluvial y est de plus en plus cultivé, mais celui-ci est particulièrement sensible à la pyriculariose, une maladie causée par un champignon microscopique, qui peut entraîner de graves pertes de production.

Mon stage s'inscrit dans le cadre du projet de recherche GARP (Gestion Agronomique de la Résistance du riz à la Pyriculariose) initié en 2010 pour vérifier l'hypothèse selon laquelle l'apport d'azote diminuerait la capacité de résistance naturelle à la pyriculariose. Ce projet a été financé par l'Agence nationale de la recherche (ANR).

Mon travail au sein du CIRAD consistait à montrer à l'aide du modèle linéaire généralisé, l'action conjointe des facteurs azote et système de culture sur l'intensité de la maladie, et en vérifiant si cette action passe ou non par le statut azoté des plantes ou par la densité du couvert végétal. Les différentes études statistiques ont été faites à l'aide du logiciel SAS, et l'extraction des données à partir du langage SQL.

ABSTRACT

From April 2nd till July 31st, 2013, I made my internship in the CIRAD (Center of international cooperation in agronomic research for the development) in Montpellier, within the Persyst department (Performances of the tropical systems of production and transformation). During this internship I worked on one of the main pluvial rice's disease in Madagascar: the pyriculariose.

Indeed in Madagascar, rice is the basic food of the population. Pluvial rice is increasingly cultivated there, but this one is particularly sensitive to the pyriculariose, a disease caused by a microscopic fungus, which can bring fairly high losses of production.

My internship is a part of the GARP research project (Agronomic Management of the rice's Resistance in Pyriculariose) initiated in 2010 to verify hypothesis the contribution of nitrogen would decrease the pyriculariose natural resistance capacity. This project is funded by the French Agence nationale de la recherche (ANR).

My work within the CIRAD consisted in showing with the generalized linear model, the joint action of nitrogen and farming system factors on the intensity of the disease, and to check if this action finds its path through by the nitrogenous status of plants or through the density of the plant place setting. The various statistical studies were made with the software SAS and the extraction of the data from the language SQL.

REMERCIEMENTS

Je tiens tout d'abord à remercier mon maître de stage Eric GOZÉ, biométricien, ainsi que Sandrine AUZOUX, informaticienne au CIRAD de Montpellier qui m'ont accueillie, aidée et conseillée durant toute la durée de mon stage.

Je remercie également Mathilde SESTER, Julie DUSSERE et Héloïse HERAUD qui, depuis Madagascar, ont répondu à toutes mes questions, et ont conçu, adapté et rempli la base de données qui m'a servi durant le stage.

Pour finir je remercie l'Institut Universitaire Professionnalisé, département Statistique et Informatique Décisionnelle l'Université de Paul Sabatier pour m'avoir permis de réaliser ce stage.

SOMMAIRE

RÉSUMÉ	III
ABSTRACT	IV
REMERCIEMENTS	V
HISTORIQUE DES MODIFICATIONS	1
1. INTRODUCTION	2
2. PRÉSENTATION DU CIRAD	3
3. DOCUMENTS APPLICABLES, DE RÉFÉRENCE ET TERMINOLOGIE	5
3.1 DOCUMENTS APPLICABLES	5
3.2 DOCUMENTS DE REFERENCE	5
3.3 TERMINOLOGIE	6
4. ORGANISATION DU DÉVELOPPEMENT	7
4.1 PRESENTATION DE L'UNITÉ DE RECHERCHE	7
4.2 ORGANISATION DU TRAVAIL / PLANIFICATION	7
5. DÉMARCHE DE DÉVELOPPEMENT	8
5.1 DÉFINITION DU SUJET	8
5.2 MÉTHODES ET OUTILS UTILISÉS	10
5.2.1 MÉTHODES	10
5.2.2 OUTILS	14
5.3 PROCESSUS DE DÉVELOPPEMENT	15
5.3.1 LE DISPOSITIF EXPERIMENTAL	15
5.3.2 LA BASE DE DONNEES	18
5.3.3 ANALYSE STATISTIQUE	24
6. CONCLUSION	48
7. ANNEXES	49
7.1 REQUETES SQL	49
7.1.1 DONNEES D'INFESTATION DE LA PYRICULARIOSE SUR LES PANICULES .	49
7.1.2 DONNEES D'INFESTATION DE LA PYRICULARIOSE SUR LES FEUILLES.....	49
7.2 MACROS SAS	50
7.2.1 MACRO 1 : IMPORTATION DES DONNEES	50
7.2.2 MACRO 2 : PRELIMINAIRES	51
7.2.3 MACRO 3 : ANALYSE	54
7.2.4 MACRO 4 : APPEL DES DIFFERENTES MACROS	55
7.3 SORTIES SAS	56
7.3.2 ANDRANO PYRICULARIOSE FOLIAIRE 2011-2012	56

HISTORIQUE DES MODIFICATIONS

Version	Date	Modification
V1.0	25/04/2013	Création du document
V1.1	02/05/2013	Amélioration des parties 2, 4.1, 5.1 et 5.2
V1.2	13/05/2013	Ajout du paragraphe 5.3.2
V1.3	21/05/2013	Ajout de la partie 5.3.3
V1.4	15/06/2013	Ajout des parties 6 et 7
V1.5	28/07/2013	Mise en forme et finalisation

Tableau 1: Historique des modifications

1. INTRODUCTION

A Madagascar, le riz est l'aliment de base de la population. Dans la région des Hautes Terres, la saturation des terroirs irrigables pousse au développement du riz pluvial, particulièrement sensible à la pyriculariose, une maladie causée par un champignon microscopique, qui peut entraîner de graves pertes de production.

La sensibilité du riz à cette maladie dépend de la variété utilisée, et elle est augmentée par l'utilisation intensive d'engrais azotés. Une hypothèse serait que l'apport ponctuel d'azote détournerait la plante de son activité de résistance naturelle à la pyriculariose. Un système de culture utilisant des plantes auxiliaires couvrant le sol permettrait alors une meilleure régulation de l'offre d'azote et une meilleure expression de la résistance à la maladie.

C'est dans ce cadre-là que le projet GARP a été mis en place, en partenariat avec le FOFIFA (Centre National de la Recherche Appliquée au Développement Rural de Madagascar), le CIRAD et l'université de Madagascar.

L'objectif de mes 4 mois de stage au sein du Centre de coopération internationale en recherche agronomique pour le développement (CIRAD) de Montpellier, est de construire une interprétation des résultats de ces essais en montrant l'action conjointe des facteurs azote et système de culture sur l'intensité de la maladie, et en vérifiant si cette action passe ou non par le statut azoté des plantes ou par la densité du couvert végétal. Les techniques statistiques mises en œuvre sont celles du modèle linéaire généralisé, avec des variables explicatives quantitatives (les facteurs étudiés) et qualitatives (le statut azoté de la plante et sa densité) ; elles permettent de choisir entre différents modèles d'explication possible de l'action de ces facteurs. Les techniques informatiques sont l'extraction de données à l'aide du langage SQL, et la programmation de macros sur le logiciel SAS.

2. PRÉSENTATION DU CIRAD

Le CIRAD (Centre de coopération internationale en recherche agronomique pour le développement) est un centre de recherche français qui répond, avec les pays du Sud, aux enjeux internationaux de l'agriculture et du développement.

- **Mission**

En partenariat avec les pays du Sud dans leur diversité, le CIRAD produit et transmet de nouvelles connaissances, pour accompagner leur développement agricole et contribuer au débat sur les grands enjeux mondiaux de l'agronomie. Organisme de recherche finalisée, le CIRAD établit sa programmation à partir des besoins du développement, du terrain au laboratoire, du local au planétaire.

- **Activités**

Ses activités relèvent des sciences du vivant, des sciences sociales et des sciences de l'ingénieur appliquées à l'agriculture, à l'alimentation et aux territoires ruraux.

Le CIRAD s'engage au plus près des hommes et de la Terre sur des défis complexes et évolutifs : sécurité alimentaire, intensification écologique, maladies émergentes, devenir des agricultures des pays du Sud.

- **Stratégie scientifique**

Le CIRAD concentre ses recherches autour de 6 axes scientifiques prioritaires. Il est présent en priorité dans le cadre de dispositifs de recherche en partenariat (14 dans le monde et 7 dans l'outre-mer français).

- **Partenariats**

Le CIRAD dispose d'un réseau mondial de partenaires et de 12 directions régionales, à partir desquelles il mène des activités de coopération avec plus de 90 pays. Ses partenariats bilatéraux s'inscrivent dans des dynamiques multilatérales d'intérêt régional.

En France métropolitaine, il met à la disposition de la communauté scientifique nationale et internationale un important dispositif de recherche et de formation situé principalement à Montpellier.

Performances des systèmes de production et de transformation tropicaux (Persyst) :

Le département scientifique Performances des systèmes de production et de transformation tropicaux s'intéresse aux processus biologiques en interaction avec les conditions environnementales et les interventions techniques.

Ses recherches sont réalisées à l'échelle de la parcelle, du troupeau, de l'exploitation agricole ou du produit tout au long de sa filière de transformation. Elles concernent les relations entre le génotype, l'environnement et le système technique.

Organigramme :

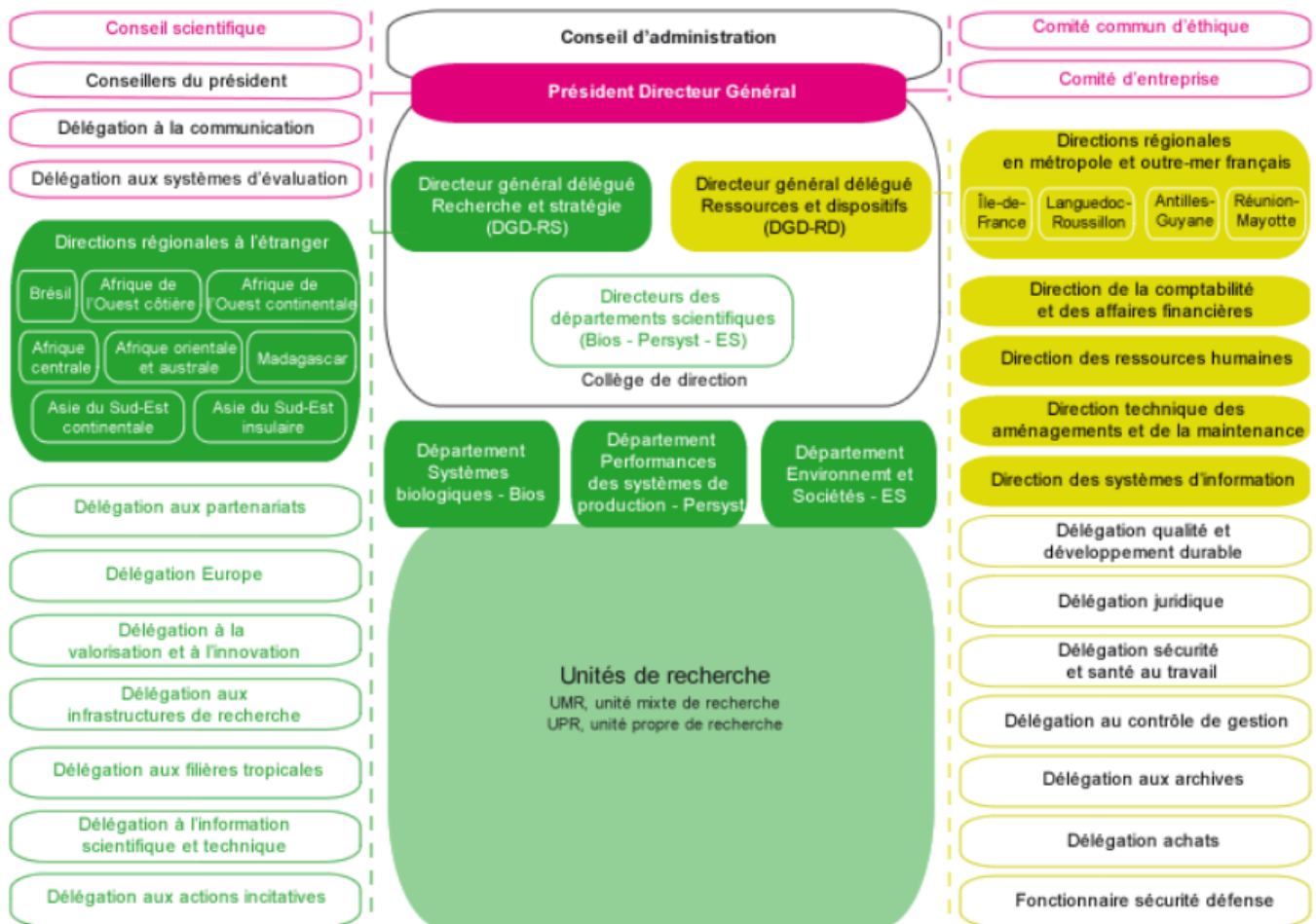


Figure 1 Organigramme du CIRAD

3. DOCUMENTS APPLICABLES, DE RÉFÉRENCE ET TERMINOLOGIE

3.1 DOCUMENTS APPLICABLES

Les documents applicables sont : le rapport de stage, les différents scripts SAS, les résultats statistiques ainsi que les requêtes SQL utilisées.

3.2 DOCUMENTS DE REFERENCE

▪ Cours

- ✓ [DR 1] Cours modèles linéaires généralisés dispensés en 2ème année M1 à l'IUP SID (Université de Paul Sabatier, Toulouse).
- ✓ [DR 2] Cours de base de données dispensés en première année L3 à l'IUP SID (Université de Paul Sabatier, Toulouse).
- ✓ [DR 3] Cours de SAS dispensés en première année L3 à l'IUP SID (Université de Paul Sabatier, Toulouse).

▪ Sites internet

- ✓ [DR 4] Introduction au modèle linéaire généralisé
http://math.agrocampus-ouest.fr/infoglue/DeliverLive/digitalAssets/43620_IntroMLG.pdf
- ✓ [DR 5] Modèles linéaires généralisés
http://maths.cnam.fr/IMG/pdf/Presentation_MODGEN_02_2007.pdf
- ✓ [DR 6] Riz (*Oryza sativa*) Connaissance des végétaux - Végétation et tallage
<http://www.afd.be/~plant-ch/riz/connais/veg-tal-riz.htm>

▪ Livres

- ✓ [DR 9] McCullagh, P, et John A Nelder. *Generalized Linear Models*. London; New York: Chapman and Hall, 1989. 532 p.
- ✓ [DR 10] William G. Cochran, et Gertrude M. Cox. *Experimental Design*. New York : John Wiley & sons, Inc 1957. 611p.
- ✓ [DR 11] Pierre Dagnelie, *Théorie et méthodes statistiques Applications agronomiques vol. 2*. Les presses agronomiques de Gembloux 1970. 463 p.
- ✓ [DR 12] Gérard. Philippeau. *Théorie des plans d'expérience. Application à l'agronomie*. Institut technique des céréales et des fourrages. 1982. 205 p.
- ✓ [DR 13] Joe Whittaker. *Graphical Models in applied multivariate statistics*. John Wiley & sons. 2009. 448 p.
- ✓ [DR 7] Sas Institute, *SAS Guide to Macro Processing*, Version 6 2nd edition, Cary NC Sas Institute Inc., 1990. 319 p.
- ✓ [DR 8] Sas Institute, *The Glimmix Procedure*, Version June 2006, Cary NC Sas Institute Inc., 2006. 255 p.

3.3 TERMINOLOGIE

CIRAD :	Centre de coopération internationale en recherche agronomique pour le développement
Biomasse :	Ensemble de la matière organique d'origine végétale ou animale
Canopé :	Couche supérieure des forêts, notamment en forêts primaires ou tropicales, soumise au rayonnement solaire, caractérisée par la présence d'un écosystème particulièrement riche de biodiversité et de productivité biologique. Par extension, en agronomie, couche supérieure d'un couvert végétal cultivé.
Dualex :	Appareil de mesure qui permet d'estimer les teneurs en chlorophylle (selon le même principe que le SPAD) et en flavonols des feuilles.
LAI :	Leaf Area Index (indice de surface foliaire). Dans un couvert végétal rapport de la surface des feuilles à la surface de sol que les plantes occupent.
Panicule :	Grappe d'épis ou d'épillets de certaines graminées. Panicules du riz
Phénologie :	Science ayant pour objet l'étude des phénomènes qui marquent la vie des plantes et des animaux au cours de l'année. Ces événements séparent des stades phénologiques. Par exemple la floraison sépare le stade végétatif du stade fructifère.
Poquet :	Trou dans lequel on dépose plusieurs graines
Pyriculariose :	Une maladie des organes aériens du riz (feuilles, tiges et panicules) causée par un champignon <i>Magnaporthe grisea</i> .
SCV :	Sous Couvert Végétal : qui consiste à semer les cultures directement sur un sol protégé par une couverture végétale vivante ou morte ; plutôt que de labourer préalablement le sol
SPAD :	Appareil portatif qui permet d'estimer la teneur en chlorophylle des feuilles
Sévérité :	Proportion moyenne des 4 feuilles terminales atteinte par la pyriculariose
Talle :	Nom féminin donné aux branches ou aux jeunes tiges qui s'élèvent, soit de la racine, soit de la tige souterraine des plantes annuelles ou herbacées. Les talles forment par leur réunion une touffe plus ou moins considérable.

4. ORGANISATION DU DÉVELOPPEMENT

4.1 PRESENTATION DE L'UNITÉ DE RECHERCHE

▪ Systemes de culture annuels

A la lumière des crises récentes qui ont touché les cours des produits agricoles, l'usage des ressources, la gestion de l'environnement, il devient essentiel d'imaginer des manières de produire, performantes certes, mais aussi plus propres, plus économes, diversifiées et résilientes, pour permettre aux paysans du Sud de vivre mieux dans le monde de demain.

Une ambition : concevoir des **systemes de production performants et durables** pour les cultures annuelles, tant du point de vue économique que social et environnemental.

Des domaines de compétences diversifiés : analyse des interactions entre génotype et environnement, bioénergies, entomologie, écophysiologie, agronomie, technologie ; appliqués à la **canne à sucre** et au **coton**. A ces compétences est venue s'ajouter celles des membres d'une équipe provenant de l'UMR System.

Cinq équipes opérationnelles :

- Conception et évaluation des systemes de culture annuels - Cesca
- Caractérisation et gestion intégrée des risques d'origine biotique - Carabe
- Élaboration et caractérisation de la qualité des produits des systemes de culture annuels - Qualité
- Télédétection, systemes d'information, techniques de simulations et analyses spatiales - Artists

- **Systemes de culture et riziculture durables – SCRiD**

4.2 ORGANISATION DU TRAVAIL / PLANIFICATION



Figure 2 Diagramme de Gantt

5. DÉMARCHE DE DÉVELOPPEMENT

5.1 DÉFINITION DU SUJET

- **Le riz pluvial à Madagascar**

A Madagascar, le riz est à la fois la culture principale et l'aliment de base de la population avec une consommation moyenne annuelle par individu supérieure à 120 Kilos. La région des Hautes Terres est densément peuplée et ses agriculteurs cultivent traditionnellement du riz irrigué.

Comme il n'y a plus de surfaces disponibles pour l'aménagement de nouvelles rizières irriguées, le développement de la riziculture pluviale sur les collines habituellement réservées aux cultures du maïs, du haricot, de la patate douce, du manioc ou au pâturage des zébus, est apparu comme une option à explorer pour augmenter la production de riz.

- **La pyriculariose**

La pyriculariose est la principale maladie du riz pluvial à Madagascar. Elle est causée par un champignon *Magnaporthe grisea*. Dans sa forme la plus virulente, le champignon attaque la tige du riz au niveau des épis et provoque sa nécrose en empêchant la circulation des éléments nutritifs.

La pyriculariose provoque fréquemment la perte de 20 % de la récolte, voire parfois sa destruction totale dans des zones où le climat est frais et humide.

La pyriculariose du riz sur feuilles et panicules est considérée comme la plus grave maladie de cette céréale majeure dont la production mondiale prévue en 2009 atteint 675 Mt.

Cette maladie, décrite dans tous les pays rizicoles, est le plus souvent liée à l'intensification de la culture suscitée par la révolution verte. Les très importants travaux de recherche menés pour en limiter les impacts n'ont pas permis d'éliminer de graves pertes de récolte. Les statistiques récentes chiffrent les pertes mondiales directement dues à cette maladie à l'équivalent de la consommation annuelle de 60 millions d'habitants avec notamment pour la Chine des pertes évaluées à 3% de la récolte, estimée à 200 Mt en 2009.

- **Projet GARP (Gestion Agronomique de la Résistance du riz à la Pyriculariose)**

Le projet GARP a été initié en 2010 pour vérifier l'hypothèse selon laquelle l'apport ponctuel d'azote détournerait la plante de son activité de résistance naturelle à la pyriculariose.

Une expérimentation a été menée dans deux lieux contrastés des Hautes Terres de Madagascar, avec 3 facteurs étudiés : la variété, l'apport d'azote et le système de culture. Les variables mesurées sont le développement et le statut azoté de la plante, le pourcentage de la surface des feuilles atteint par le champignon, le taux de grains vides à cause de la maladie, la croissance des plantes et le rendement. Les données ont été saisies depuis 2010 dans des feuilles MS-Excel et regroupées dans une base MS-Access.

Les expérimentateurs ont observé que les pratiques agronomiques sont des facteurs importants de modulation du niveau des épidémies de pyriculariose du riz. Il est notamment connu que la fertilisation azotée favorise les épidémies, alors que des systèmes « agroécologiques » qui incluent le semis direct sur couverture végétale (SCV), permettent d'atteindre des rendements élevés en limitant le niveau des épidémies. L'hypothèse qui est privilégiée est que le contrôle des épidémies résulterait de différences de la dynamique du métabolisme de l'azote. Les systèmes agroécologiques réguleraient les apports d'azote et permettraient une expression satisfaisante de la résistance partielle, alors que dans des systèmes conventionnels les apports d'azote massifs induiraient une répression des voies métaboliques impliquées dans la résistance. Des expérimentations préliminaires menées en serre avec des apports azotés différents indiquent que la voie du métabolisme de l'azote serait antagoniste de celle des mécanismes de défense.

5.2 MÉTHODES ET OUTILS UTILISÉS

5.2.1 MÉTHODES

Dans notre cas l'ANOVA n'était pas appropriée, en effet les données concernant la pyriculariose étant des pourcentages, avec l'ANOVA il y aurait eu un problème d'échelle des prédicteurs linéaires (qui peuvent sortir de l'intervalle 0 à 100), ainsi qu'un problème de variance non homogène et d'additivité des effets. C'est pourquoi nous utilisons le modèle linéaire généralisé, avec comme fonction de lien le logit et une fonction de variance non standard.

▪ **Modèle linéaire généralisé**

Les modèles linéaires généralisés (MLG) sont une généralisation du modèle linéaire dans les cas où la loi de la variable réponse fait partie de la famille exponentielle : c'est le cas de lois continues comme la loi exponentielle ou la loi gamma et de lois discrètes comme la loi binomiale ou de Poisson. Les modèles linéaires généralisés utilisés le plus souvent sont des modèles de régression logistiques pour des données binaires et des modèles log-linéaires pour des données « poissonniennes ».

Le modèle linéaire généralisé contient trois éléments fondamentaux :

- Le prédicteur linéaire : $\eta = X\beta$
- La fonction de lien : $\eta = g(\mu)$, qui doit être inversible : $\mu = g^{-1}(\eta)$
- La loi de l'observation, conditionnelle aux régresseurs

Par exemple : $Y | \mu \Xi \mathcal{B}(\mu, n)$

La variance de Y conditionnelle au prédicteur linéaire est une fonction de son espérance, cette fonction de variance est utile dans de nombreux calculs.

$$\text{Var}(Y | \mu) = V(\mu)$$

- **Régression logistique**

Dans notre cas la variable que l'on souhaite expliquer est une proportion d'infection de feuille ou de panicule infecté. Nous avons gardé la fonction de lien logit du modèle de régression logistique, adaptée à des observations binomiales.

▪ **La surdispersion**

L'ajustement d'un modèle à un jeu de données particulier peut faire apparaître des résultats (concernant les résidus, notamment) non compatibles avec les hypothèses du modèle.

Si les variances estimées à partir des résultats du modèle sont supérieures aux variances estimées, on parle de surdispersion. Celle-ci peut être due à l'omission de variables explicatives, à l'existence d'erreurs dans les variables explicatives, à une mauvaise spécification du modèle ou à la non-homogénéité de l'environnement.

On peut mettre en évidence la présence de surdispersion dans les données relativement au modèle obtenu en testant la somme des carrés des résidus réduits (ou résidus de Pearson) qui suit une loi de χ^2 . Le rapport de ce χ^2 à ses degrés de liberté permet d'estimer un paramètre de surdispersion.

En présence de surdispersion, la fonction de variance (qui exprime la variance en fonction de l'espérance) est donc inadaptée, et le modèle peut alors produire des estimations biaisées des écarts types des paramètres, et selon les cas des paramètres eux-mêmes.

Les travaux sur la méthode de la quasi-vraisemblance [Wedderburn, 1974] montrent que l'ajustement d'un modèle linéaire généralisé dépend de la fonction de variance mais ne dépend pas de la forme précise de la distribution utilisée. Une prise en compte de la surdispersion peut donc être introduite au moyen d'une modification de la fonction de variance, sans qu'il soit nécessaire de postuler une forme particulière de distribution (parmi les distributions possédant la même fonction de variance).

▪ **Quasi vraisemblance**

L'estimation par quasi-vraisemblance est un moyen de permettre la surdispersion, qui est, une plus grande variabilité dans les données que prévu dans le modèle statistique utilisé. C'est le plus souvent le cas lorsque la loi des observations est une loi Binomiale ou de Poisson.

Les paramètres des modèles de quasi-vraisemblance peuvent être estimés avec une simple extension des algorithmes utilisés pour des modèles linéaires généralisés.

Seule une relation entre la moyenne et la variance est spécifiée sous la forme d'une fonction de variance donnant la variance en fonction de la moyenne.

En général, cette fonction peut inclure un facteur multiplicatif appelé le paramètre de surdispersion qui est estimé à partir des données, comme peut l'être la variance pour un modèle gaussien.

La « quasi vraisemblance » est une fonction des paramètres évaluée aux observations, à l'instar de la vraisemblance. Cette quasi-vraisemblance est définie comme :

$$Q(\mu, y) = \int_y^\mu \frac{y - t}{\phi V(t)} dt$$

▪ **Détermination empirique de la relation moyenne variance par régression du log de la variance sur le log de la moyenne**

Dans notre cas ; la relation moyenne variance des proportions de surface foliaire atteinte par la maladie ne suit pas une loi binomiale, ce qui signifie que la relation moyenne variance n'est pas :

$$\text{▪ } V(\mu) = n \mu (1 - \mu)$$

Or dans le livre *Generalized Linear Model* de P. McCullagh et J.A. Nelder on a pu lire que, pour une autre maladie, la Rouille de l'orge la relation moyenne variance était du type :

$$\text{▪ } V(\mu) = k \mu^2 (1 - \mu)^2$$

En ce qui concerne les données sur la pyriculariose nous en sommes arrivés à la fonction suivante :

$$\text{▪ } V(\mu) = k \mu^p (1 - \mu)^p$$

Avec p variable suivant le relevé

Dans l'exemple de McCullagh et Nelder, la fonction de variance est choisie de façon à rendre la variance des résidus réduits indépendante de leur moyenne, d'après le graphique croisant les résidus et les valeurs prédites, ce qui fait intervenir un jugement subjectif reposant sur un nombre limité de points. Nous verrons plus loin que les données modélisées sont les infections moyennes par parcelle, et sont les moyennes de dix mesures par parcelle. Comme la variance intra-parcelle contribue à la variance des moyennes par parcelle, nous avons étudié la variance intra-parcelle comme une fonction de la moyenne par parcelle pour proposer une première approximation de la fonction de variance, avant de vérifier les résidus comme l'ont fait Mc Cullagh et Nelder.

Afin d'estimer p nous utilisons une méthode décrite dans *Théorie et méthodes statistiques vol.2* qui consiste à faire une régression linéaire du logarithme de la variance sur le logarithme de la moyenne, et de retenir le coefficient comme paramètre de notre fonction de variance.

▪ **Analyse de covariance « généralisée »**

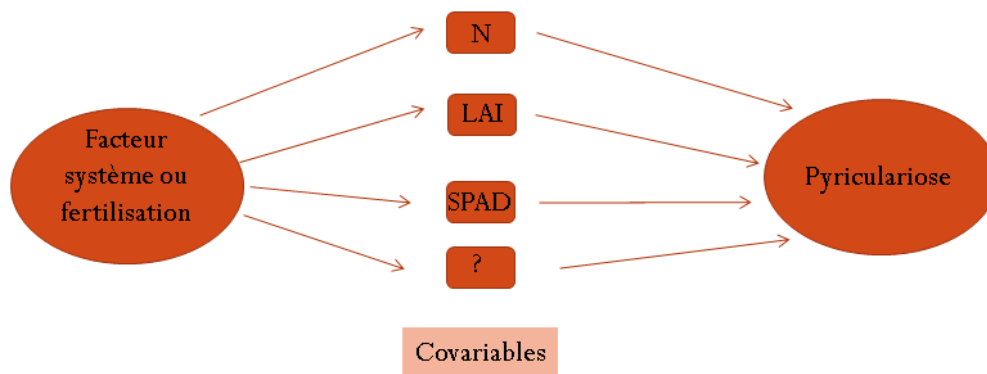
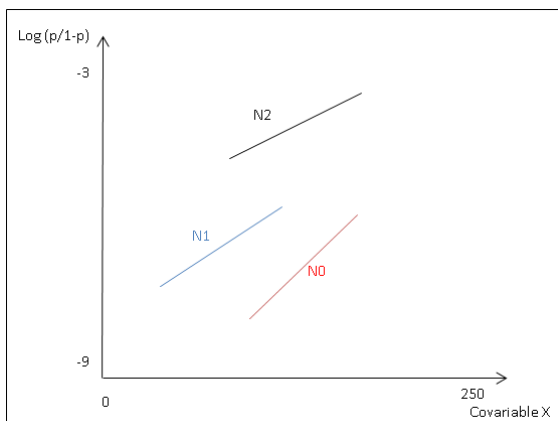


Figure 3 Modèle graphique analyse de covariance

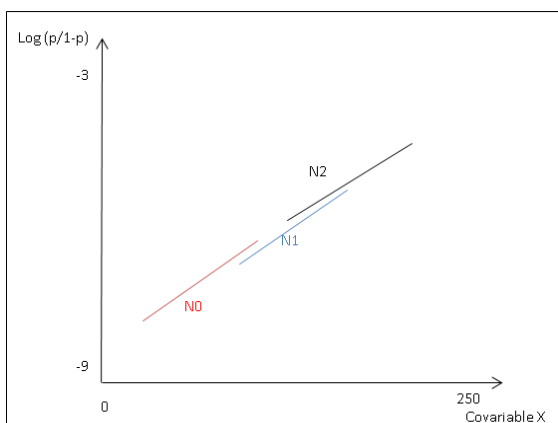
Quand il faut proposer des éléments d'explication du mode d'action des facteurs étudiés, des variables intermédiaires sont ajoutées dans un modèle graphique : le facteur agit à travers des états de la plante mesurés à l'aide de covariables. Le facteur peut aussi agir directement sans passer par les covariables, mais par un mécanisme inconnu.



Comment va-t-on décider de l'existence ou pas de flèches dans notre modèle graphique ? En utilisant la généralisation de l'analyse de covariance.

Prenons un exemple imaginaire d'une relation entre une variable X, et l'infection par la pyriculariose.

Dans la figure du haut, cette relation va être modulée par l'application d'azote : suivant la dose appliquée, la droite de régression va être décalée. L'on peut donc affirmer qu'il y a un effet de l'azote qui ne passe pas par la variable X.



Au contraire, dans la figure du bas, les droites de régression sont alignées : tout se passe comme si l'effet de l'azote sur la pyriculariose passait par la covariable X.

Figure 4 Illustrations exemples analyse de covariance

Tout au long mon stage j'ai été amenée à utiliser les outils suivants :

- **Microsoft Access**

Microsoft Access (officiellement Microsoft Office Access) est un SGBD relationnel édité par Microsoft. Il fait partie de la suite bureautique MS Office Pro.

MS Access est composé de plusieurs programmes : le moteur de base de données Microsoft Jet, un éditeur graphique, une interface pour manipuler les bases de données, et le langage de programmation Visual Basic for Applications.

- **SAS**

SAS, leader mondial des solutions de Business Analytics, est le plus important éditeur indépendant du marché du décisionnel. Avec plus de 60 000 sites clients répartis dans 135 pays. SAS est un logiciel complet qui couvre une large gamme des méthodes d'analyse en Statistique. De conception américaine, il est développé par la société SAS Institute Inc. basée à Cary, en Caroline du Nord; et a acquis, depuis sa mise sur le marché en 1976, une situation dominante dans de nombreuses branches d'activités économiques. Le système SAS se présente sous la forme d'un ensemble de modules logiciels adaptés pour la gestion et l'analyse statistique de gros volumes de données (tableaux de plusieurs giga-octets) et la création de rapports de synthèse. (Source : Wikipedia)

- **Proc Glimmix**

La procédure GLIMMIX permet d'adapter des modèles linéaires généralisés où des effets aléatoires gaussiens peuvent affecter le prédicteur linéaire. On connaît ces modèles sous le nom de Modèles Linéaires Généralisés Mixtes

Dans notre cas, l'atout principal de cette procédure est qu'elle permet de spécifier un MLG où la loi conditionnelle des observations n'est connue que par sa fonction de variance.

5.3 PROCESSUS DE DÉVELOPPEMENT

5.3.1 LE DISPOSITIF EXPERIMENTAL

Les variables :

Deux sites sont représentés dans la base de données : Andranomanelatra et Ivory comportant chacun 144 parcelles.

- La fertilisation

Toutes les parcelles reçoivent tous les ans une fumure organique à base de fumier. En plus de cette fumure de base, elles reçoivent ou non une fertilisation minérale uniquement azotée.

En ce qui concerne la fertilisation minérale, 3 modalités sont utilisées par site et par campagne. Les modalités de fertilisation ne sont pas les mêmes sur les deux sites.

F1 est la même modalité dans toutes les situations : pas d'apport minéral d'azote. F2 est un apport fractionné de l'azote en 3 fois ; aux doses 30, 23 et 23 unités d'azote à Andranomanelatra et 30, 15 et 15 unités d'azote à Ivory. Fv est un apport fractionné à dose variable suivant les lieux et les années (voir tableau 2)

		Andranomanelatra	Ivory
Précédemment		SCV	Friche
2009-2010	F1	sans azote	Labour Complet
	F2	30N + 23N + 23N	
	Fv	30N + 23N	
2010-2011	F1	Attaque de vers blancs	sans azote
	F2		30N + 15N + 15N
	Fv		60N + 30N + 30N
2011-2012	F1	sans azote	sans azote
	F2	30N + 23N + 23N	30N + 15N + 15N
	Fv	60N + 46N + 46N	60N + 30N + 30N
2012-2013	F1	Attaque de vers blancs	sans azote
	F2		30N + 15N + 15N
	Fv		30N + 30N

Tableau 2: Récapitulatif des fertilisations

- **Les variétés**

Trois variétés de riz sont semées dans chacun des sites afin de représenter une gamme de sensibilité à la pyriculariose. La variété tolérante ou résistante est utilisée comme témoin afin de voir quels sont les rendements obtenus pour une variété qui n'est pas (ou très peu) touchée par la Pyriculariose mais qui subit néanmoins les mêmes conditions que les deux autres variétés. Les variétés de riz semées ont été choisies pour leur adaptation aux conditions climatiques du site où elles sont utilisées.

		Andrano		Ivory	
2009-2010	v1	FOFIFA 154	très sensible	Labour Complet	
	v2	FOFIFA 161	moyennement sensible		
	v3	FOFIFA 172	résistante		
2010-2011	v1	Attaque de vers blancs		FOFIFA 154	très sensible
	v2			B 22	sensible
	v3			Nerica 4	tolérante
2011-2012	v1	FOFIFA 154	très sensible	FOFIFA 154	très sensible
	v2	FOFIFA 161	moyennement sensible	B 22	sensible
	v3	FOFIFA 172	résistante	Nerica 4	tolérante
2012-2013	v1	Attaque de vers blancs		FOFIFA 154	très sensible
	v2			B 22	sensible
	v3			Nerica 4	tolérante

Tableau 3 Récapitulatif des variétés

- **Les systèmes de culture**

Sur les différentes parcelles deux types de système de culture sont présents :

- Le labour qui est un travail de la terre qui consiste à l'ouvrir et à la retourner en profondeur avant de semer la culture
- Le SCV (sous couvert végétal) qui consiste à semer les cultures directement sur un sol protégé par une couverture végétale vivante ou morte.

- **Les blocs**

Sur ces deux sites, le dispositif expérimental est un dispositif en blocs. Chaque site est constitué de quatre répétitions formant des blocs d'un seul tenant.

Le dispositif est factoriel, ce qui signifie que les 18 traitements possibles sont placés de façon aléatoire dans les 18 parcelles présentes sur chaque bloc. Les parcelles élémentaires ont une surface de 42m². Les infections de pyriculariose sont relevées de manière répétée sur 10 poquets par parcelle sur une diagonale choisie avant chaque comptage (en rouge sur la figure 6)

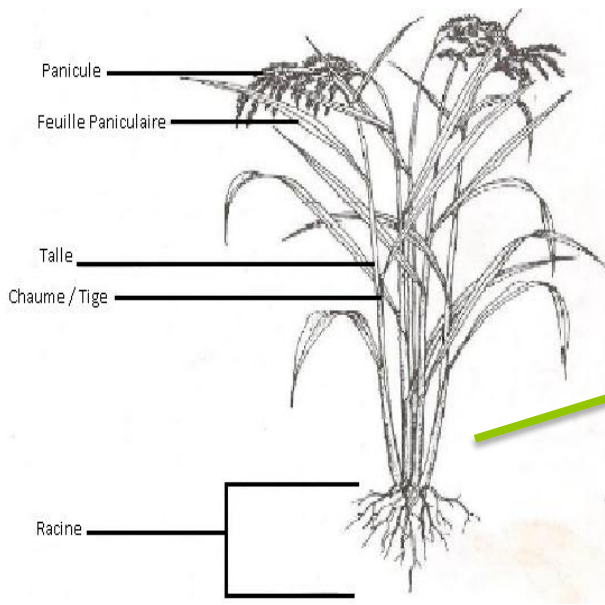


Figure 5 Schéma anatomique d'un plant de riz

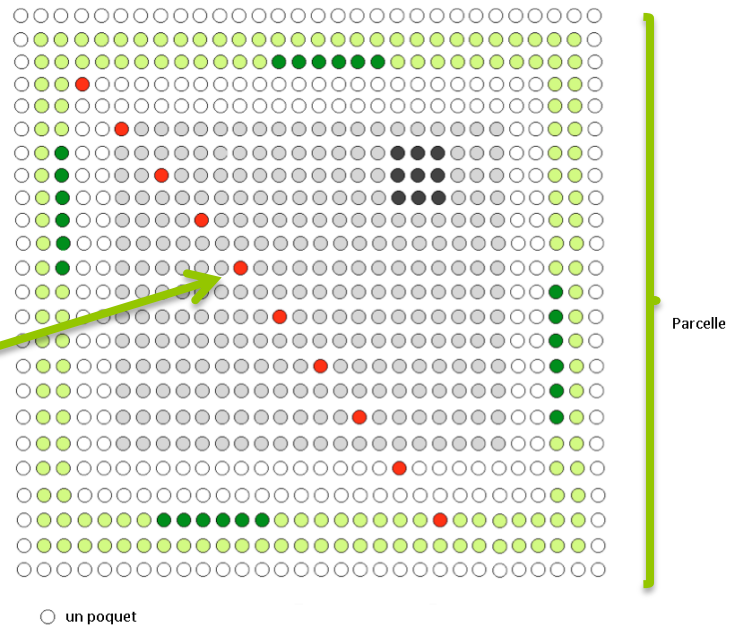


Figure 6 Schéma d'une parcelle expérimentale

5.3.2 LA BASE DE DONNEES

A partir du dispositif expérimental décrit précédemment, a été créée une base de données qui a été mise à ma disposition. Elle a été conçue dans le but de recenser et stocker l'ensemble des données récoltées dans le cadre du projet sur les différents sites, et dans un second temps pour extraire des données afin de réaliser des traitements statistiques.

Modèle Conceptuel des Données (MCD) :

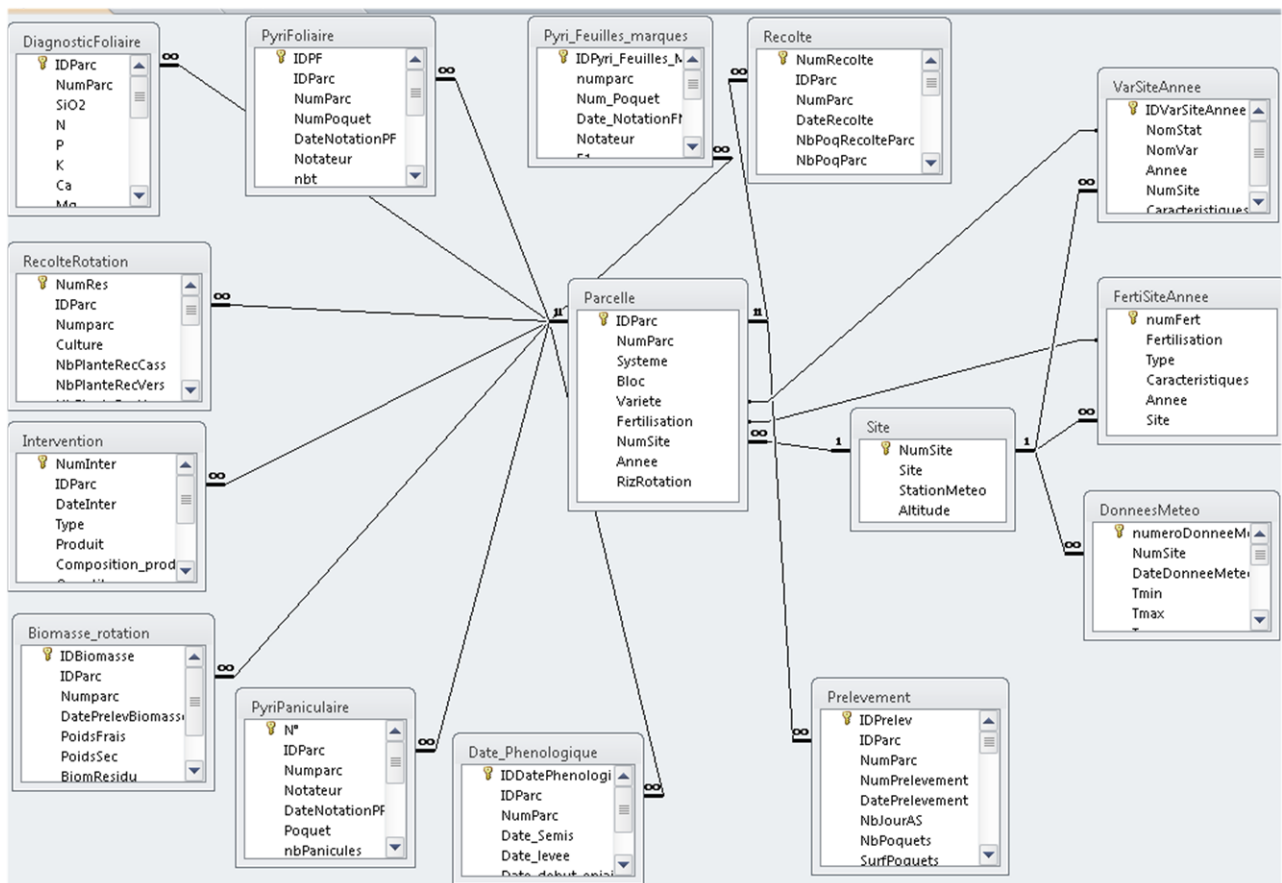


Figure 7 Modèle conceptuel de données

Le travail effectué au long de mon stage ne concerne que quelques tables dont la description détaillée se trouve ci-dessous :

- Parcelle

Nom de la variable	Description
IDParc	Identifiant de la parcelle
NumParc	Numéro de la parcelle
Système	Système de culture utilisé (Labour ou SCV)
Bloc	Bloc auquel appartient la parcelle (A, B, C ou D)
Variete	Variété de riz plantée sur la parcelle (v1, v2 ou v3)
Fertilisation	Fertilisation utilisée sur la parcelle (F1, F2 ou Fv)
NumSite	Numéro du site auquel appartient la parcelle (1 ou 2)
Annee	Année de la campagne
RizRotation	1 si riz planté sur la parcelle, 0 sinon

Tableau 4 Description de la table Parcelle

- Site

Nom de la variable	Description
NumSite	Identifiant du site
Site	Nom du site (Ivory ou Andranomanelatra)
StationMeteo	Nom de la station météo présente sur le site
Altitude	Altitude à laquelle se trouve le site

Tableau 5 Description de la table Site

- PyriFoliaire (une ligne par talle)

Nom de la variable	Description
IDPF	Identifiant de la notation
IDParc	Identifiant de la parcelle
NumParc	Numéro de la parcelle
NumPoquet	Numéro du poquet noté appartenant à la parcelle
DateNotationPF	Date à laquelle les mesures ont été prises
Notateur	Nom de la personne qui a pris les mesures
Nbt	Nombre de talles sur le poquet
Nbti	Nombre de talles infectées sur le poquet
Fn	% d'infection de la 1 ^{re} feuille
Fn-1	% d'infection de la 2 ^e feuille
Fn-2	% d'infection de la 3 ^e feuille
Fn-3	% d'infection de la 4 ^e feuille
Fn-4	% d'infection de la 5 ^e feuille
Sev1234	Sévérité de l'infection moyenne sur les feuilles 1 à 4 $((fn+fn-1+fn-2+fn-3)/4)$

Tableau 6 Description de la table PyriFoliaire

- PyriPaniculaire (une ligne par poquet)

Nom de la variable	Description
N°	Identifiant de la notation
<i>IDParc</i>	Identifiant de la parcelle
<i>NumParc</i>	Numéro de la parcelle
Notateur	Nom de la personne qui a pris les mesures
DateNotationPP	Date à laquelle les mesures ont été prises
Poquet	Numéro du poquet noté appartenant à la parcelle
nbPanicules	Nombre de panicules sur le poquet
nbPaniculesInfect	Nombre de panicules infectées sur le poquet
P1	% d'infection de la panicule 1
P2	% d'infection de la panicule 2
P3	% d'infection de la panicule 3
P4	% d'infection de la panicule 4
P5	% d'infection de la panicule 5
Sev	Incidence de l'infection sur le poquet (nbPaniculesInfect/nbPanicules)
SevPoq	Sévérité de l'infection au poquet (Moyenne(P1,P2,P3,P4,P5) *nbPaniculesInfect /nbPanicules)

Tableau 7 Description de la table PyriPaniculaire

- FertiSiteAnnee

Nom de la variable	Description
numFert	Identification de la fertilisation
Fertilisation	Nom de la fertilisation
Type	Type de fertilisation
Caractéristiques	Caractéristique de la fertilisation
Année	Année au cours de laquelle la fertilisation a été utilisée
Site	Numéro du site sur lequel la fertilisation a été utilisée

Tableau 8 Description de la table FertiSiteAnnee

- Prélèvement

Nom de la variable	Description
IDPrelev	Identifiant du prélèvement
<i>IDParc</i>	Identifiant de la parcelle
<i>NumParc</i>	Numéro de parcelle
NumPrelevement	Numéro du prélèvement (1, 2, 3 ou 4)
DatePrelevement	Date à laquelle le prélèvement a été effectué
NbJourAS	Date du prélèvement en jours après semis
NbPoquets	Nombre de poquets
SurfPoquets	surface des poquets prélevés
SPAD	Spad dernière feuille
DuaChl	Dualex Chlorophylle dernière feuille
DuaFlav	Dualex Flavonoïde dernière feuille
DuaNBI	Dualex indicateur dernière feuille
HautCano	Hauteur de la Canopée
HautTot	Hauteur du poquet en étirant en cm
LAI SunScan	Leaf Area Index en m ² /m ²
NbPlants	Nombre de plantes récoltées sur 8 poquets
NbPlantsParm2	Nombre de plantes par m ²
NbTalles	Nombre de talles
NbTallesParPlt	Nombre de talles par plante
NbTallesParm2	Nombre de talles par m ²
BiomPlanteg	Biomasse aérienne des plantes de la placette prélevée en g
BiomPlanteTHa	Biomasse aérienne des plantes de la placette prélevée ramené en tonne par hectare
NPlante	teneur en azote de plante entière

QteNPlante	Quantité d'azote absorbé par les plantes entières en kg par ha
CourbeRefN	Données de la courbe de référence de dilution de l'azote pour le riz
INN	Indice de nutrition calculé d'après la courbe de référence
NNH4Plante	teneur en azote soluble sous forme ammonium ppm de plante entière
QteNNH4Plante	Quantité N-NH4 absorbé par les plantes entières en kg/ha
NNO3Plante	teneur en azote soluble sous forme nitrate ppm de plante entière
QteNNO3Plante	Quantité N-NO3 absorbé par les plantes entières en kg/ha
BiomFeuilles	Biomasse des feuilles paniculaires prélevées en g
NFeuilles	teneur en azote ppm dernière feuille ligulée
QteNFeuilles	Quantité d'azote absorbé par dernière feuille ligulée
NNH4Feuilles	teneur en azote soluble sous forme ammonium ppm dernière feuille ligulée
QteNNH4Feuilles	Quantité d'azote soluble sous forme ammonium ppm absorbé par dernière feuille ligulée
NNO3Feuilles	teneur en azote soluble sous forme nitrate ppm dernière feuille ligulée
QteNNO3Feuilles	Quantité d'azote soluble sous forme nitrate ppm absorbé par dernière feuille ligulée

Tableau 9 Description de la table Prélèvement

Les données se relevées par talle ou par poquet, mais les traitements ont été appliqués à des parcelles ; ce sont donc les parcelles les unités expérimentales, et les analyses statistiques se feront sur les moyennes parcellaires. Afin de pouvoir effectuer ces analyses statistiques il nous faut calculer au niveau parcellaire l'infection moyenne par la pyriculariose, appelée sévérité. Pour cela, il faut créer plusieurs vues depuis le logiciel Access :

- Données d'infestation de la Pyriculariose sur les feuilles

Dans un premier temps, les calculs de sévérité sont effectués à la talle (cf. Annexe 7.1.2). Au niveau du poquet deux sévérités moyennes différentes sont calculées, une pondérée et une autre non pondérée :

- Pour une feuille : l'infection d'une feuille est le pourcentage de surface foliaire attaqué par la maladie.
- Pour une talle : la moyenne non pondérée est la moyenne des infections des seules feuilles infectées. La moyenne pondérée est la moyenne non pondérée multipliée par la proportion de feuilles infectées (elle est tout simplement l'infection moyenne sur toutes les feuilles infectées ou non)

Dans un second temps, les moyennes sont calculées mais cette fois ci, au niveau de la parcelle, toujours en gardant la moyenne pondérée et celle qui ne l'est pas.

Voici un extrait de la vue Sev_Parcelle_Foliaire2 :

IDPArc	NumParc	Bloc	Systeme	Fertilisation	Variete	Annee	NumSite	DateNotationPF	Incidence	Severite	SeveritePor
450	A48	C	Labour	Fv	v1	2009-2010	1	11/02/2010	55,991	1,048	0,60
420	A20	B	SCV	Fv	v1	2009-2010	1	11/02/2010	6,402	0,725	0,13
455	A52	C	SCV	Fv	v1	2009-2010	1	11/02/2010	7,814	0,436	0,0
464	A60	D	SCV	Fv	v1	2009-2010	1	11/02/2010	10,583	0,761	0,09
496	A9	A	SCV	Fv	v1	2009-2010	1	11/02/2010	12,268	0,692	0,08
477	A72	D	Labour	F1	v2	2009-2010	1	26/01/2010	1,825	0,2	0,00
375	A11	A	SCV	F1	v2	2009-2010	1	26/01/2010	0,556	0,05	0,00
466	A62	D	SCV	F1	v2	2009-2010	1	26/01/2010	1,576	0,117	0,0
375	A11	A	SCV	F1	v2	2009-2010	1	11/02/2010	3,5	0,125	0,01
426	A26	B	SCV	F1	v2	2009-2010	1	11/02/2010	1,458	0,275	0,01

Figure 8 Extrait vue "Sev_Parcelle_Foliaire2"

- Données d'infestation de la Pyriculariose sur les panicules

Tout d'abord la sévérité est calculée par poquet afin de s'assurer de l'exactitude des calculs, puis celle-ci est liée avec les données parcellaires (Bloc, Système, Fertilisation ...). (cf. Annexe 7.1.1)

Ensuite la sévérité est calculée à la parcelle en faisant la moyenne des sévérités pour chaque poquet présent dans la parcelle, et l'incidence moyenne par parcelle est calculée en faisant la moyenne des incidences au poquet.

Voici un extrait de la vue Sev_Parcelle_Paniculaire2 :

IDPArc	NumParc	Bloc	Systeme	Fertilisation	Variete	NumSite	DateNotationPP	Incidence	Severite	logMeanSev	logVarSev2
507	I1	A	Labour	F1	v1	2	24/02/2010	41,04	7,011	1,94748	3,8204
508	I10	A	Labour	Fv	v1	2	24/02/2010	47,32	11,203	2,41618	5,1707
541	I13	A	Labour	F2	v1	2	24/02/2010	79,48	30,757	3,42612	5,5359
552	I14	A	Labour	Fv	v3	2	24/02/2010	0	0	0	
559	I16	A	Labour	F2	v3	2	24/02/2010	0	0	0	
561	I18	A	Labour	F1	v2	2	24/02/2010	16,81	0	0	
564	I20	B	Labour	F1	v3	2	24/02/2010	0	0	0	
565	I21	B	Labour	F2	v2	2	24/02/2010	26,24	5,562	1,71596	3,6101
566	I22	B	Labour	F2	v3	2	24/02/2010	0	0	0	
567	I23	B	Labour	F1	v1	2	24/02/2010	30,41	3,834	1,34391	4,0654

Figure 9 Extrait vue "Sev_Parcelle_Paniculaire2"

Pour des besoins purement esthétiques les résultats sont arrondis, ceux qui seront utilisés pour les analyses ne le seront pas.

Ces vues sont alors exportées vers Excel, pour pouvoir être importées sous SAS et ensuite être analysées.

5.3.3 ANALYSE STATISTIQUE

5.3.3.1 MACROS SAS.

Les analyses faites précédemment ont été faites à l'aide de macros SAS pour l'année 2011-2012. Mon travail a consisté à reprendre ses macros et à les adapter afin de pouvoir faire les analyses sur toutes les données qui étaient à notre disposition.

Durant mon stage j'ai été amenée à créer des macros différentes pour les données sur la Pyriculariose Foliaire et Paniculaire. Voici l'exemple des macros créées pour les données sur la Pyriculariose Foliaire

La première sert à importer les données sous SAS (cf. Annexe 7.2.1) :

Pour faire fonctionner cette macro, il faut dans un premier temps affecter les 2 macros variables « lieu » et « date » afin de pouvoir importer les données correctement.

Une fois les macros variables affectées, la macro permet d'importer le fichier Excel provenant de la base de données, à l'aide d'une PROC IMPORT.

Dans un second temps une nouvelle table SAS est créée, avec les variables que l'on souhaite garder et en y intégrant la variable « proportion ». En effet, ce sont des pourcentages qui ont été stockés, mais le modèle linéaire généralisé concerne ici des proportions.

Puis enfin à l'aide des instructions « %if » et « %else » la macro exécute une PROC SQL pour ne retenir que les données dont nous auront besoin pour les analyses.

Une seconde macro « préliminaire » sert à faire, comme son nom l'indique les analyses préliminaires (cf. Annexe 7.2.2) :

Cette macro produit dans un premier temps le graphique du logarithme de la variance et de celui de la moyenne, ceci à l'aide de la procédure « GPLOT ». Puis une PROC EXPORT est exécutée ensuite pour pouvoir avoir les données en fichier texte et si besoin les utiliser sous R. Une procédure « TABULATE » nous permet par la suite d'obtenir un tableau des sévérités moyennes

Dans un second temps afin de tester la fonction de variance, plusieurs PROC GLIMMIX sont exécutées les unes à la suite des autres, ainsi qu'une régression de la valeur absolue du résidu de Pearson sur la valeur prédite".

On utilise ici le Test de Glejser : prévu pour le modèle linéaire, ce test propose de régresser la valeur absolue des résidus de la régression sur la valeur prédite, ici sur le prédicteur linéaire pour le modèle linéaire généralisé

Ceci nous permettra d'évaluer et de proposer une fonction de variance en étudiant la variabilité intra-parcelle, puis à la valider en vérifiant la relation entre dispersion des résidus de Pearson et la valeur prédite.

La macro « modèle feuille » (cf. Annexe 7.2.3) qui est lancée après la Macro « préliminaire », permet de faire l'analyse statistique avec la fonction de variance retenue précédemment, et pouvoir ajouter une covariable si on le souhaite grâce à la macro variable « covar ».

Avant l'exécution de cette macro il faut attribuer à la macro variable « b » le coefficient que l'on souhaite attribuer à notre fonction de variance, suivant les résultats de la macro « préliminaire ».

Cette macro permet dans un premier temps d'avoir un tableau de la proportion moyenne selon les variables « bloc », « fertilisation », « système » et « variété ». La covariable peut être ajoutée dans ce tableau si elle est présente.

Enfin la procédure GLIMMIX est exécutée avec ou sans covariable et avec le nouveau coefficient pour la fonction de variance.

Un script SAS a été ensuite créé afin d'appeler toutes les macros précédentes (cf. Annexe 7.2.4) :

On peut voir, en premier lieu que les macros variables sont affectées au lieu « Ivory » et à la date « 18JAN2013 ». Puis les instructions « %include » vont chercher et exécuter les fichiers où sont stockées les différentes macros. L'instruction %preliminaire, par exemple appelle la macro « préliminaire ».

Enfin suivant le cas, on peut être amené à ne sélectionner que certaines variétés ou certaines fertilisations ayant trop de données nulles car cela empêche l'algorithme de converger. En effet une proportion à 0 donne un logit qui tend vers l'infini, ce qui pose problème. Dans ce cas-là après l'importation des données, une PROC SQL permettant de ne garder que les données pertinentes est exécutée.

5.3.3.2 PYRICULARIOSE

Les études statistiques concernant la pyriculariose, ont été faites sur la pyriculariose foliaire et paniculaire, sur les différentes années. Toutes les études ne pouvant pas être mises dans le rapport de stage, nous présenterons ci-dessous quelques exemples des différentes situations possibles.

EXEMPLES

- **Ivroy Pyriculariose Foliaire 2011-2012**

L'utilisation de la première macro SAS nous permet d'obtenir la régression du log de la variance sur le log de la moyenne l'équation suivante :

$$\text{Regression Equation:} \\ \log VarSev = -0.479086 + 1.55024 * \log MeanSev$$

Ce qui nous permet d'en déduire la fonction de variance suivante :

$$_mu_ **1.55024 * (1-_mu_)**1.55024$$

Nous obtenons les résultats suivants :

Model Information	
Data Set	PYRIFOL.OBSPARCELLE3
Response Variable	proportion
Response Distribution	Unknown
Link Function	Logit
Variance Function	$_mu_ **1.55024 * (1-_mu_)**1.55024$
Variance Matrix	Diagonal
Estimation Technique	Quasi-Likelihood
Degrees of Freedom Method	Residual

Tableau 10 Information sur le modèle

Class Level Information		
Class	Levels	Values
Bloc	4	A B C D
Fertilisation	3	F1 F2 Fv
Systeme	2	Labour SCV
Variete	2	v1 v2

Tableau 11 Information sur les niveaux des classes

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	32	6.33	0.0017
Systeme	1	32	56.90	<.0001
Fertilisation	2	32	13.46	<.0001
Fertilisatio*Systeme	2	32	0.49	0.6175
Variete	1	32	4.58	0.0401
Systeme*Variete	1	32	3.33	0.0776
Fertilisatio*Variete	2	32	0.33	0.7219
Fertil*System*Variet	2	32	0.77	0.4719

Tableau 12 Type III

Fit Statistics	
-2 Log Quasi-Likelihood	1.03
Quasi-AIC (smaller is better)	31.03
Quasi-AICC (smaller is better)	46.51
Quasi-BIC (smaller is better)	58.78
Quasi-CAIC (smaller is better)	73.78
Quasi-HQIC (smaller is better)	41.47
Pearson Chi-Square	0.69
Pearson Chi-Square / DF	0.02

Tableau 13 : Fit Statistics

La fonction de variance est proposée à une constante k près (Pearson Chi-Square / DF), ce qu'on spécifie avec l'instruction random residual. On remarque qu'elle est très faible (0.02). En effet nous n'avons pas spécifié d'effectif n dans la fonction, puisque le pourcentage est un pourcentage de surface et non d'effectif, il est peut-être assez peu variable.

Le tableau « Type III Tests of Fixed Effects » nous indique la présence d'un effet bloc, d'un effet variété. Mais également d'un effet système et fertilisation hautement significatif avec des p -value inférieures à 0.0001.

Nous ne remarquons également aucune interaction entre les différents facteurs.

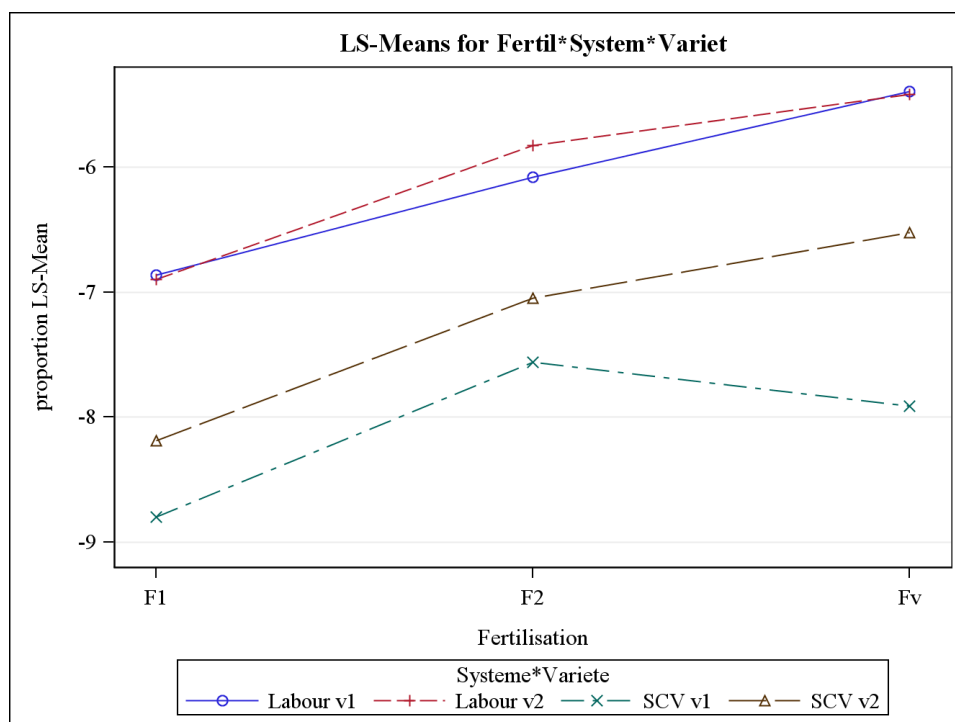


Figure 10 Evolution du logit de la proportion suivant la variété la fertilisation et le système de culture

Le graphique ci-dessus nous permet de voir que le système de culture en Labour (courbe rouge et bleue) est plus touché que le système de culture en SCV (courbe verte et marron). Rappelons que nous sommes en présence d'une faible attaque de pyriculariose. (entre $\log(p/1-p) = -9$ et $\log(p/1-p) = -5$)

Nous pouvons voir également que la variété 2 (courbe rouge et marron) est légèrement plus touchée par la pyriculariose que la variété 1 (courbe bleue et verte). Ce qui ne semble pas logique car la variété 1 est censée être plus sensible à la maladie que la variété 2, mais après discussion avec les spécialistes il s'avère que la notion de variété sensible ou moyennement sensible est variable suivant les différentes souches de pyriculariose qui peuvent être présentes une année ou une autre.

Enfin nous remarquons que généralement, plus la fertilisation azotée augmente plus l'infection augmente.

- Andrano Pyriculariose Foliaire 2011-2012

L'utilisation de la première macro SAS nous permet d'obtenir la régression du log de la variance sur le log de la moyenne l'équation suivante :

<p>Regression Equation: $\log\text{VarSev} = -0.388887 + 1.555385 * \log\text{MeanSev}$</p>

Cela nous permet d'en déduire la fonction de variance suivante :

$_{\mu} ** 1.555385 * (1 - _{\mu}) ** 1.555385$

Nous obtenons les résultats suivants :

Model Information	
Data Set	PYRIFOL.OBSPARCELLE3
Response Variable	proportion
Response Distribution	Unknown
Link Function	Logit
Variance Function	$_{\mu} ** 1.555385 * (1 - _{\mu}) ** 1.555385$
Variance Matrix	Diagonal
Estimation Technique	Quasi-Likelihood
Degrees of Freedom Method	Residual

Tableau 14 Information sur le modèle

Class Level Information		
Class	Levels	Values
Bloc	4	A B C D
Fertilisation	3	F1 F2 Fv
Systeme	2	Labour SCV
Variete	2	v1 v2

Tableau 15 Information sur les niveaux des classes

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	80	1.85	0.1451
Systeme	1	80	41.49	<.0001
Fertilisation	2	80	4.28	0.0172
Fertilisatio*Systeme	2	80	1.12	0.3310
Variete	1	80	8.45	0.0047
Systeme*Variete	1	80	8.78	0.0040
Fertilisatio*Variete	2	80	0.40	0.6713
Fertil*System*Variet	2	80	0.79	0.4578

Tableau 16 Type III

Fit Statistics	
-2 Log Quasi-Likelihood	5.37
Quasi-AIC (smaller is better)	35.37
Quasi-AICC (smaller is better)	41.44
Quasi-BIC (smaller is better)	73.67
Quasi-CAIC (smaller is better)	88.67
Quasi-HQIC (smaller is better)	50.85
Pearson Chi-Square	5.58
Pearson Chi-Square / DF	0.07

Tableau 17 : Fit Statistics

Le tableau « Type III Tests of Fixed Effects » nous indique la présence d'un effet variété et d'un effet fertilisation. Mais également d'un effet système hautement significatif avec une p-value inférieure à 0.0001.

Nous remarquons également une interaction entre les facteurs système et variété significative.

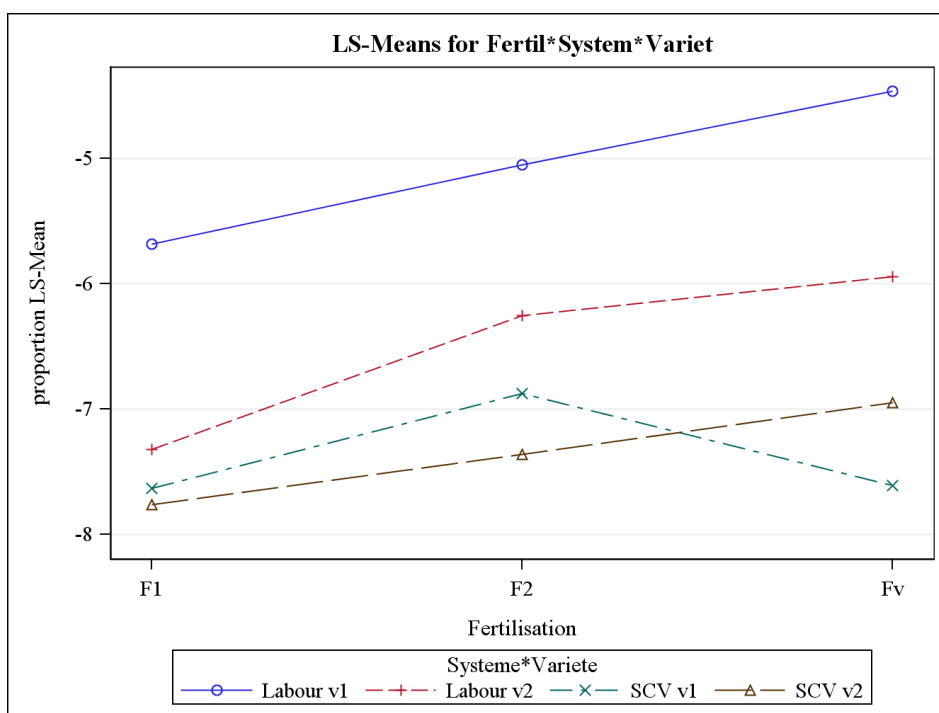


Figure 11 Evolution du logit de la proportion suivant la variété la fertilisation et le système de culture

Le graphique ci-dessus nous permet de voir que le système de culture en Labour (courbes rouge et bleue) est plus touché que le système de culture en SCV (courbes verte et marron). Il apparaît que la variété 1 (courbes bleue et verte) est légèrement plus touchée par la pyriculariose que la variété 2 (courbes rouge et marron), sauf dans le cas de la fertilisation Fv où la variable v1 est moins touchée que v2 en système de culture SCV.

Toutefois cette particularité n'a pas lieu d'être interprétée, car il n'y a pas d'interaction significative entre fertilisation et variété (P=0.6713), le croisement des courbes peut donc être fortuit.

Enfin nous pouvons voir que généralement, plus la fertilisation azotée augmente plus l'infection augmente. Sauf toujours dans le cas de la fertilisation Fv pour la variable où l'infection chute.

- Andrano Pyriculariose Foliaire 2009-2010

L'utilisation de la première macro SAS nous permet d'obtenir la régression du log de la variance sur le log de la moyenne l'équation suivante :

<p>Regression Equation $\log\text{VarSev} = -0.015873 + 1.608212 * \log\text{MeanSev}$</p>
--

Cela nous permet d'en déduire la fonction de variance suivante :

$_{\mu} ** 1.608212 * (1 - _{\mu}) ** 1.608212$

Nous obtenons les résultats suivants :

Model Information	
Data Set	PYRIFOL.OBSPARCELLE3
Response Variable	proportion
Response Distribution	Unknown
Link Function	Logit
Variance Function	$_{\mu} ** 1.608212 * (1 - _{\mu}) ** 1.608212$
Variance Matrix	Diagonal
Estimation Technique	Quasi-Likelihood
Degrees of Freedom Method	Residual

Tableau 18 Information sur le modèle

Class Level Information		
Class	Levels	Values
Bloc	4	A B C D
Fertilisation	3	F1 F2 Fv
Systeme	2	Labour SCV
Variete	2	v1 v2

Tableau 19 Information sur les niveaux des classes

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	33	3.87	0.0178
Systeme	1	33	1.36	0.2512
Fertilisation	2	33	0.79	0.4616
Fertilisatio*Systeme	2	33	1.54	0.2302
Variete	1	33	92.34	<.0001
Systeme*Variete	1	33	4.12	0.0505
Fertilisatio*Variete	2	33	0.85	0.4383
Fertil*System*Variet	2	33	3.45	0.0436

Tableau 20 Type III

Fit Statistics	
-2 Log Quasi-Likelihood	2.85
Quasi-AIC (smaller is better)	32.85
Quasi-AICC (smaller is better)	47.85
Quasi-BIC (smaller is better)	60.91
Quasi-CAIC (smaller is better)	75.91
Quasi-HQIC (smaller is better)	43.45
Pearson Chi-Square	1.75
Pearson Chi-Square / DF	0.05

Tableau 21 : Fit Statistics

Le tableau « Type III Tests of Fixed Effects » nous indique la présence d'un effet variété et d'un effet bloc.

Il apparait également une interaction entre les facteurs système, variété et fertilisation significative. Ce sera le seul cas où nous trouverons une interaction triple significative au cours du stage.

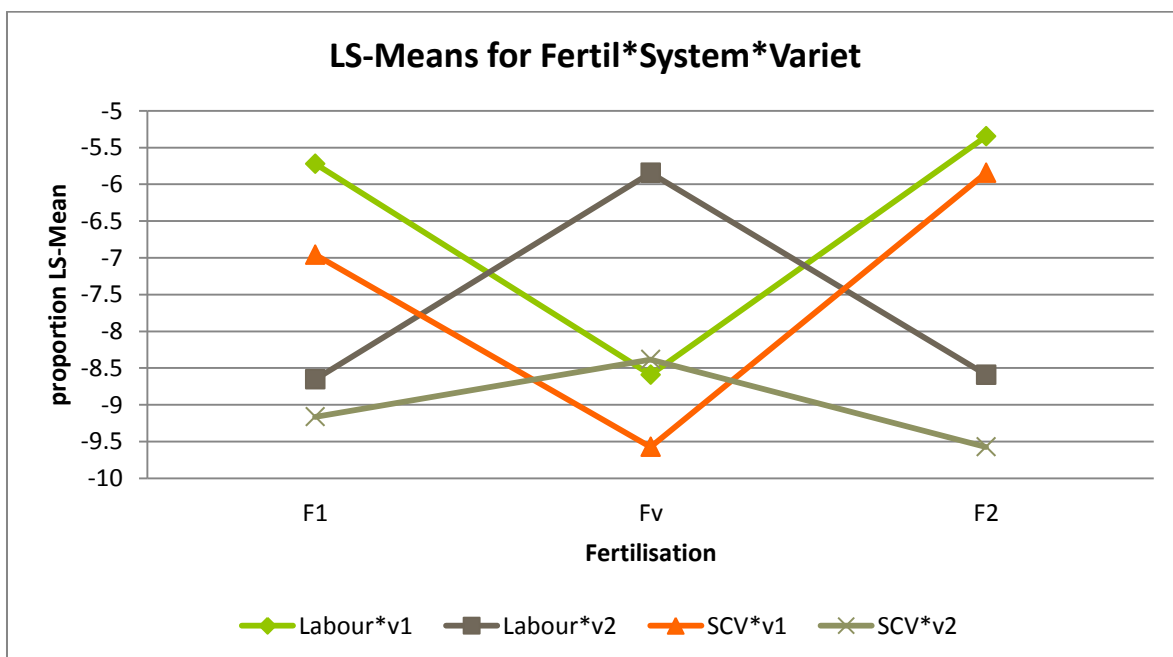


Figure 12 Evolution du logit de la proportion suivant la variété la fertilisation et le système de culture

Le graphique ci-dessus est très difficile à interpréter en effet, on peut voir les variétés ont chacune un même comportement quel que soit le type de système de culture. Cependant les tests disent le contraire : interaction ferti x système x variété juste significative; et interaction système x variété juste non significative.

Nous remarquons quand même que le labour est un peu plus touché par la pyriculariose que le SCV même si le facteur système de culture n'est pas ressorti comme significatif de l'analyse.

Concernant les fertilisations on remarque une infection similaire pour les deux espèces en F1 et F2 pour chaque système de culture. Seule la fertilisation variable (Fv) pose problème car soit pour la variété 1 l'infection atteint son minimum, soit pour la variété 2 elle atteint son maximum.

Il faut rappeler que nous sommes en présence d'une infection extrêmement faible et donc que les conclusions sont d'autant plus difficiles à en tirer. En effet d'une part les tests sont asymptotiques, pas forcément fiables pour des infections aussi faibles. D'autre part des niveaux aussi faibles sont certainement le résultat d'infections primaires, qui n'ont pas encore effectué de cycle de reproduction dans le plant, et sont donc plus affectées par le hasard de la répartition des parcelles que par l'effet de l'engrais ou du système.

RECAPITULATIF :

Pour avoir une vue d'ensemble sur le travail réalisé sur l'infection de la pyriculariose tout au long de mon stage, un tableau récapitulatif des résultats des différents effets a été produit.

Site	Année	Pyriculariose	Facteurs							
			Système	Fertilisation	Variété	Bloc	Syst*Ferti	Syst*Var	Ferti*Var	Syst*Var*Ferti
Andrano	2009 - 2010	Foliaire			***	*				*
		Paniculaire	**							
	2011 - 2012	Foliaire	***		***			***		
		Paniculaire	***		***	***				
Ivory	2010 - 2011	Foliaire	***	**	**	***				
		Paniculaire		**						
	2011 - 2012	Foliaire	***	***	*	**				
		Paniculaire				**				
	2012 - 2013	Foliaire		**	***	***			**	
		Paniculaire	***	**						

Tableau 22 Récapitulatif des différents effets

P-value *** <0,001, **<0,01, *<0,05

Le tableau récapitulatif ci-dessus nous indique que contrairement aux hypothèses énoncées par le projet GARP aucune interaction entre le système et la fertilisation n'a été trouvée que ce soit sur Ivory ou Andrano tout au long des 4 années d'expérimentation.

Un autre fait marquant a été mis en évidence, il s'agit de l'absence totale d'effet significatif de la fertilisation sur le site d'Andrano. Nous pouvons voir aussi le peu d'interaction entre facteurs significatif.

Concernant les différents effets, les conclusions sont les suivantes :

- Effet système de culture : Les différentes analyses ont pu montrer que le système de culture en Labour était plus touché par la pyriculariose que le système de culture sous couvert végétal (SCV), hormis en 2012-2013 où c'est l'effet inverse qui a été constaté. Une hypothèse formulée l'an dernier est que la moindre pyriculariose observée sur SCV est due à la moindre densité de végétation obtenue avec ce système. Cette hypothèse se confirme avec cette année une plus grande végétation sur SCV, accompagnée de d'avantage de pyriculariose que sur labour.
- Effet fertilisation : Nous avons pu voir que, généralement le pourcentage touché par la pyriculariose augmente avec la fertilisation. Donc l'azote aurait un effet aggravant sur la pyriculariose.
- Effet variétal: Nous avons pu montrer que globalement la variété 1 dite très sensible à la maladie est plus touchée que la variété 2 qui est censée être moyennement sensible. Il apparait que cet effet s'inversait dans quelques cas, en effet la notion de variété sensible peut être remise en cause par l'arrivée de nouvelles souches de Pyriculariose qui ne sont pas évitées par les mêmes mécanismes de résistance que les souches qui les précédaient.

5.3.3.3 ANALYSE DE COVARIANCE GENERALISEE

Si l'effet des facteurs imposés dans le modèle (mode de gestion du sol et fertilisation azotée) passe par une covariable, l'ajout de la covariable dans le modèle linéaire généralisé utilisé précédemment doit annuler l'effet des facteurs conditionnel à cette covariable, c'est-à-dire l'effet supplémentaire du facteur une fois la covariable déjà présente dans le modèle. Ce sont les effets conditionnels qui sont testés avec les tests de type III de Sas.

Nous avons choisi d'ajouter tout d'abord des covariables reflétant la densité du couvert, ainsi que des covariables reflétant le statut azoté afin de voir si l'effet du système de culture ou l'effet de la fertilisation azotée disparaît.

Les covariables dont nous disposons sont les suivantes :

- Celles contenues dans la table prélèvement : INN, LAI, Azote Foliaire, Biomasse, Dualex, SPAD, Hauteur, Azote
- Le diagnostic foliaire : SiO₂, N, P, K, Ca, Mg, Na, Cu, Mn, Zn, B, S, NNH₄, NNO₃, NAlphaAmi

Chaque covariable a été mesurée entre 2 et 4 fois selon l'année de culture. Ce qui signifie qu'une étude a dû être mise en place pour chaque prélèvement de chaque covariable et ce pour chaque site, année et type de pyriculariose (foliaire ou paniculaire). Dans un souci d'espace, seuls quelques exemples d'analyse seront détaillés dans la suite de ce document.

Les graphiques qui accompagnent ces exemples sont des graphiques avec des régressions, cela est en accord avec un modèle linéaire mais pas avec un modèle linéaire généralisé. Il faut donc les prendre pour des illustrations mais les positions des droites ne sont pas exactement celles qui seraient estimées avec un modèle linéaire généralisé.

EXEMPLES :

- **Ivory – Pyriculariose Foliaire – 2011-2012**

Lors des précédentes analyses sur la pyriculariose, les effets système de culture et fertilisation avaient été trouvés hautement significatifs (tableau de gauche). En ajoutant la covariable SPAD au prélèvement 3 dans notre modèle on remarque que l'effet fertilisation précédemment hautement significatif a été supplanté par l'ajout de la covariable SPAD.

Nous remarquons également que la covariable SPAD en elle-même n'a pas d'effet en plus de celui de la fertilisation.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	32	6.33	0.0017
Systeme	1	32	56.90	<.0001
Fertilisation	2	32	13.46	<.0001
Fertilisatio*Systeme	2	32	0.49	0.6175
Variete	1	32	4.58	0.0401
Systeme*Variete	1	32	3.33	0.0776
Fertilisatio*Variete	2	32	0.33	0.7219
Fertil*System*Variet	2	32	0.77	0.4719

Tableau 23 Type III sans covariable

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
SPADP3	1	31	1.98	0.1688
Bloc	3	31	6.08	0.0022
Systeme	1	31	11.18	0.0022
Fertilisation	2	31	2.42	0.1059
Fertilisatio*Systeme	2	31	0.56	0.5766
Variete	1	31	6.11	0.0191
Systeme*Variete	1	31	2.35	0.1352
Fertilisatio*Variete	2	31	0.47	0.6309
Fertil*System*Variet	2	31	0.88	0.4260

Tableau 24 Type III avec covariable

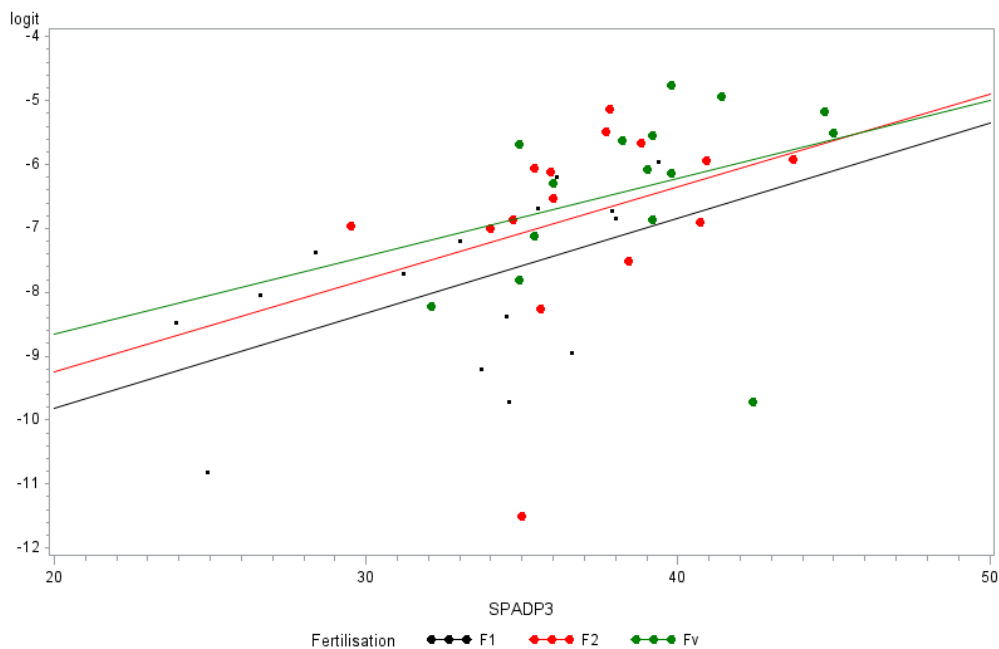


Figure 13 Evolution du logit de la proportion suivant le prélèvement 3 de la covariable SPAD

Le graphique ci-dessus, nous montre bien que les droites de régression de chaque fertilisation ont à peu près la même pente et aussi la même ordonnée à l'origine. Nous voyons que les parcelles ayant une fertilisation F1 ont une valeur SPAD moins élevée que les parcelles ayant une fertilisation F2 ou Fv.

Il apparaît enfin que plus la valeur SPAD est élevée plus la proportion de feuilles touchées par la maladie augmente.

- **Andranomanelatra– Pyriculariose Paniculaire – 2009-2010**

Lors des précédentes analyses sur la pyriculariose, seul l'effet système de culture avait été trouvé significatif (tableau de gauche). En ajoutant la covariable INN au prélèvement 1 dans notre modèle on remarque que l'effet système de culture précédemment significatif a été supplanté par l'ajout de la covariable INN.

Nous remarquons également que la covariable INN en elle-même a un effet significatif sur la maladie.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	15	1.39	0.2846
Systeme	1	15	10.52	0.0055
Fertilisation	2	15	3.06	0.0767
Systeme*Fertilisatio	2	15	0.83	0.4551

Tableau 25 Type III sans covariable

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	14	0.77	0.5320
Systeme	1	14	1.52	0.2375
Fertilisation	2	14	7.59	0.0558
Systeme*Fertilisatio	2	14	1.79	0.2037
INN1	1	14	8.07	0.0131

Tableau 26 Type III avec covariable

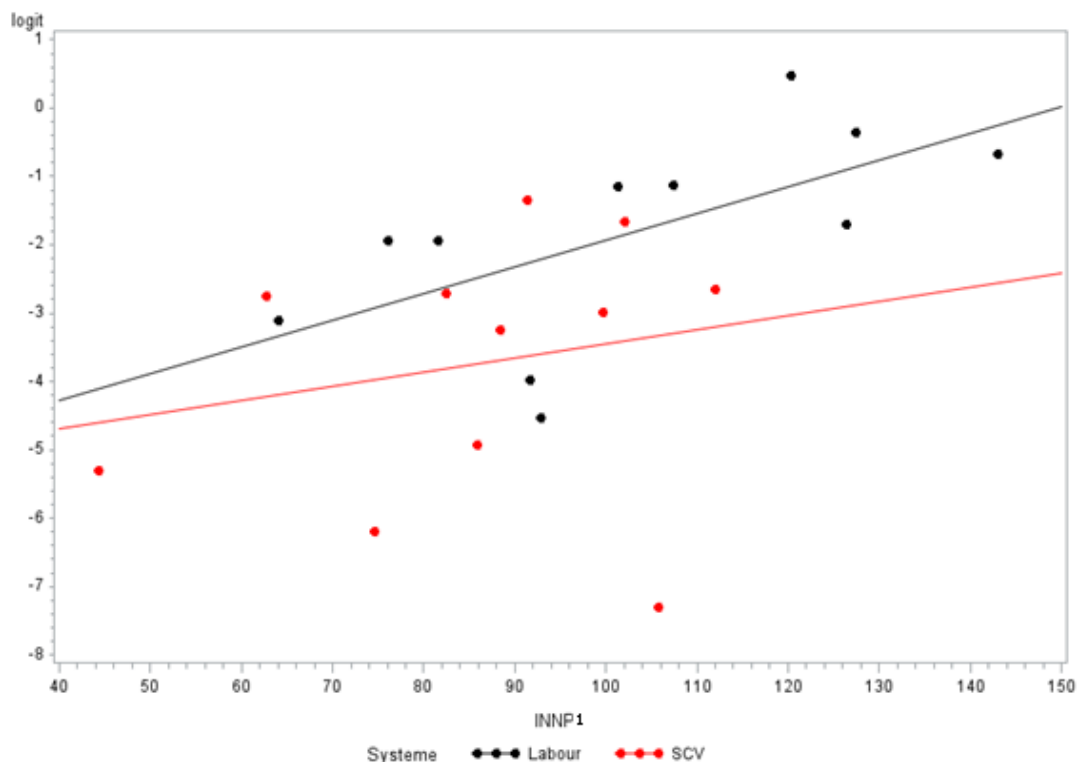


Figure 14 Evolution du logit de la proportion suivant le prélèvement 1 de la covariable INN

Le graphique ci-dessus, nous montre que les parcelles cultivées avec un système de culture en SCV ont tendance à avoir une valeur INN plus faible que les parcelles cultivées en labour. Nous voyons enfin que plus la valeur INN est élevée plus la proportion de feuilles touchées par la maladie augmente.

On voit aussi qu'il reste une grande variabilité d'INN au sein du labour comme du SCV, donc des variations d'INN qui ne sont pas causées par le système de culture. Cela explique pourquoi l'effet d'INN reste significatif même si l'effet système fait aussi partie du modèle.

- Ivory – Pyriculariose Foliaire – 2010-2011

Lors des précédentes analyses sur la pyriculariose, les effets système de culture et fertilisation avaient été trouvés significatifs (tableau de gauche). En ajoutant la covariable Dualex aux prélèvements 1 et 2 dans notre modèle on remarque que pour le prélèvement 1 comme pour le prélèvement 2 les deux effets (système de culture et fertilisation) ont été supplantés.

Nous remarquons également que les covariables Dualex aux prélèvements 1 et 2 n'ont pas d'effet significatif en elles-mêmes sur la maladie. Cela veut dire qu'il n'a pas ou pas assez de variabilité autre que celle causée par l'apport d'engrais

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	2	10	19.91	0.0003
Systeme	1	10	8.09	0.0174
Fertilisation	2	10	4.19	0.0477
Fertilisatio*Systeme	2	10	2.04	0.1804

Tableau 27 Type III sans covariable

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
DualexP2	1	9	0.30	0.5992
Bloc	2	9	12.97	0.0022
Systeme	1	9	4.89	0.0544
Fertilisation	2	9	3.39	0.0799
Fertilisatio*Systeme	2	9	1.03	0.3942

Tableau 28 Type III avec covariable

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
DualexP1	1	9	0.39	0.5490
Bloc	2	9	17.61	0.0008
Systeme	1	9	0.52	0.4872
Fertilisation	2	9	1.43	0.2893
Fertilisatio*Systeme	2	9	2.46	0.1403

Tableau 29 Type III avec covariable

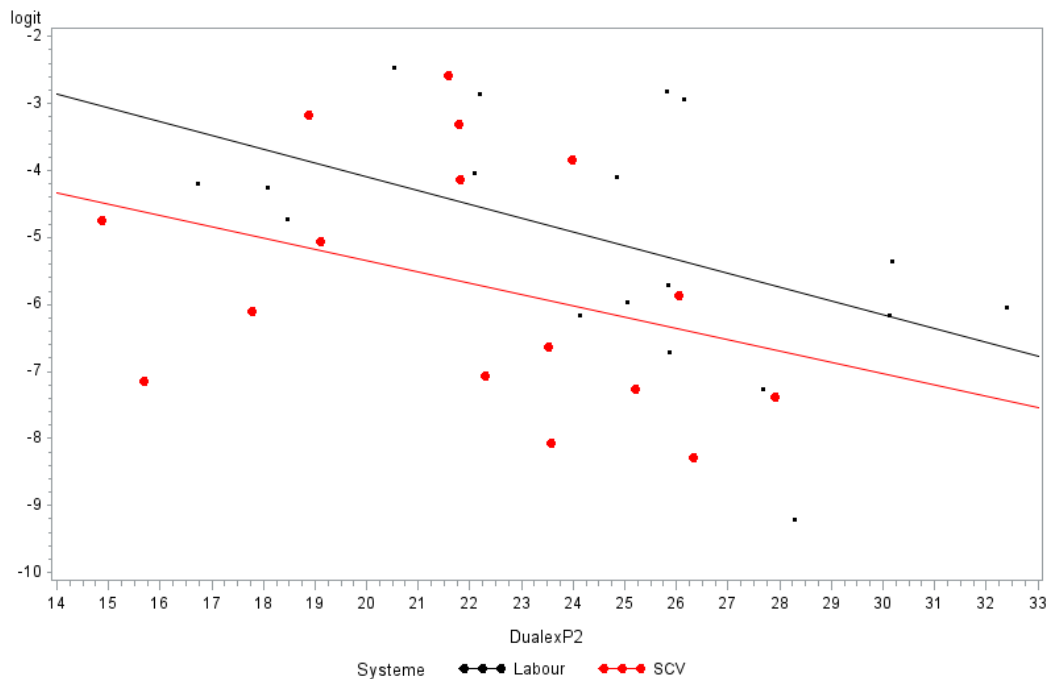


Figure 15 Evolution du logit de la proportion suivant le prélèvement 2 de la covariable Dualex

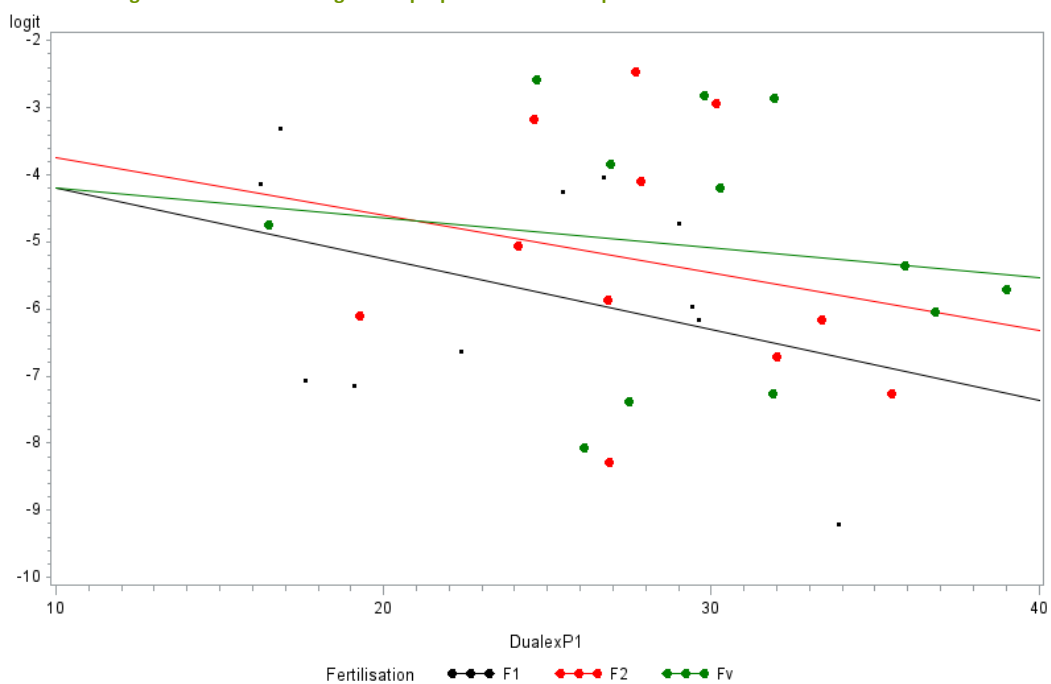


Figure 16 Evolution du logit de la proportion suivant le prélèvement 1 de la covariable Dualex

Les graphiques ci-dessus, nous montrent que les droites de régression de chaque fertilisation ont plus ou moins la même pente. Nous remarquons que les parcelles cultivées en Labour ont une valeur Dualex légèrement plus élevée que celles cultivées en labour. Ce qui saute aux yeux sur ces deux graphiques est le fait que plus la valeur Dualex augmente plus la proportion de feuilles touchées par la pyriculariose diminue, ce qui sous-entend que plus il y a d'azote plus la pyriculariose est faible. Cette interprétation va à l'encontre des hypothèses de départ. L'année 2010-2011 pour le site d'Ivory est la seule année où l'on peut tirer ce genre de conclusion.

- Ivory – Pyriculariose Foliaire – 2012-2013

Lors des précédentes analyses sur la pyriculariose, seul l'effet fertilisation a été trouvé significatif (tableau de gauche). En ajoutant la covariable LAI au prélèvement 3 dans notre modèle nous remarquons que l'effet fertilisation trouvé précédemment est supplanté par l'ajout de la covariable. En revanche il reste plusieurs interactions. Il y a donc bien un effet fertilisation, mais il dépend du système et de la variété (interaction triple et les deux interactions double sont significatives). L'action du système et de la variété sur la pyriculariose ne passent donc pas entièrement par LAI au prélèvement 3, effectivement la majeure partie des variations de LAI sont engendrées par les facteurs étudiés.

Nous voyons également que la covariable LAI au prélèvement 3 n'a pas d'effet significatif en elle-même sur la maladie: les plus grandes variations de LAI sont dues aux traitements.

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	33	14.99	<.0001
Systeme	1	33	0.07	0.7972
Fertilisation	2	33	5.43	0.0091
Fertilisatio*Systeme	2	33	2.84	0.0726
Variete	1	33	126.13	<.0001
Systeme*Variete	1	33	1.55	0.2213
Fertilisatio*Variete	2	33	7.01	0.0029
Fertil*System*Variet	2	33	2.45	0.1020

Tableau 30 Type III sans covariable

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
LAIP3	1	30	2.68	0.1122
Bloc	3	30	14.65	<.0001
Systeme	1	30	0.02	0.9024
Fertilisation	2	30	1.06	0.3582
Fertilisatio*Systeme	2	30	5.11	0.0123
Variete	1	30	50.53	<.0001
Systeme*Variete	1	30	3.02	0.0927
Fertilisatio*Variete	2	30	9.77	0.0005
Fertil*System*Variet	2	30	3.65	0.0382

Tableau 31 Type III avec covariable

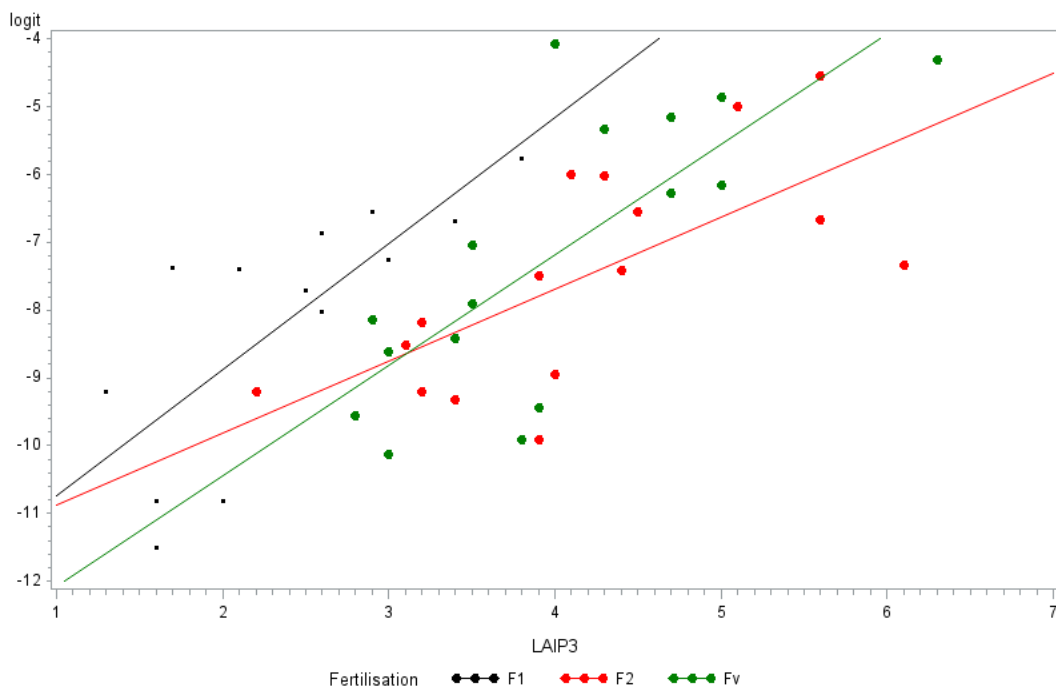


Figure 17 Evolution du logit de la proportion suivant le prélèvement 3 de la covariable LAI

Le graphique ci-dessus, nous montre que les droites de régression ont à peu près la même pente et ordonnée à l'origine et ce quelle que soit la fertilisation. Il est important de remarquer que les parcelles ayant une fertilisation F1 ont une valeur LAI moins élevée que les parcelles ayant une fertilisation Fv et F2.

Il apparaît enfin que plus la valeur LAI est élevée plus la proportion de feuilles touchées par la maladie augmente.

RECAPITULATIF :

Pour avoir une vue d'ensemble sur le travail réalisé sur les covariables tout au long de mon stage, plusieurs tableaux récapitulatifs ont été produits : deux tableaux pour les covariables de prélèvement (un pour la pyriculariose foliaire et un second pour la pyriculariose paniculaire) et deux tableaux pour les covariables de diagnostic foliaire (un pour la pyriculariose foliaire et un second pour la pyriculariose paniculaire).

✓ Covariables de la table prélèvement

Pyriculariose Foliaire											
	Andrano				Ivory						
	2009-2010		2011-2012		2010-2011		2011-2012		2012-2013		
	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	
INN	Pas d'effet système	Pas d'effet fertilisation		Pas d'effet fertilisation	P1 P2	P1 P2			Pas d'effet système		
LAI			P2 P3 P4				P3			P2 P3	
Azote Feuille			P1		P1	P1 P2		P3			
Biomasse					P1					P1	
Dualex						P1 P2	P1 P2				
SPAD						P1				P3	
Hauteur					P3	P2				P3	
Azote Plante						P1 P2	P1 P2				

Tableau 32 Récapitulatif covariables de la table prélèvement, pyriculariose foliaire

Pour la pyriculariose foliaire on peut voir sur le tableau ci-dessus que pour le site Andranomanelatra en 2011-2012 seul l'effet système de culture avait été trouvé significatif. A l'aide de l'analyse de covariance « généralisée » nous avons pu voir que cet effet système de culture était supplanté par les covariables LAI aux prélèvements 2,3 et 4, Azote Foliaire au prélèvement 1 ainsi que la Hauteur au prélèvement 3.

Concernant le site Ivory en 2010-2011 nous remarquons que toutes les covariables permettent de supplanter les 2 effets. Rien n'a permis d'expliquer cela.

En 2011-2012 nous voyons que l'effet système de culture qui avait été trouvé significatif précédemment n'est expliqué par aucune des covariables, alors que l'effet fertilisation lui est supplanté par les prélèvements 3 des covariables LAI, Azote Foliaire, SPAD et Hauteur. Enfin en 2012-2013 seul l'effet fertilisation avait été trouvé significatif lors des analyses précédentes. Les covariables LAI pour les prélèvements 2 et 3 et la biomasse pour le prélèvement 1 permettent de supplanter son effet.

Pyriculariose Paniculaire										
	Andrano				Ivory					
	2009-2010		2011-2012		2010-2011		2011-2012		2012-2013	
	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation
INN	P1	Pas d'effet fertilisation		Pas d'effet fertilisation	Pas d'effet système		Pas d'effet système	Pas d'effet fertilisation		
LAI										
Azote Feuille										
Biomasse										
Dualex										
SPAD										
Hauteur										
Azote Plante										

Tableau 33 Récapitulatif covariables de la table prélèvement, pyriculariose paniculaire

Pour la pyriculariose paniculaire, le tableau ci-dessus nous montre pour le site Andranomanelatra qu'en 2009-2010 seul l'effet système avait été trouvé significatif précédemment ; il a pu être supplanté uniquement par le premier prélèvement de la covariable INN.

Pour l'année 2011-2012 aucune covariable n'a permis d'expliquer l'effet système de culture qui avait été trouvé significatif lors des analyses précédentes.

Concernant le site Ivory, en 2010-2011 tout comme en 2012-2013 aucune des covariables n'a permis d'expliquer les différents effets trouvés significatifs précédemment .

✓ Covariables de diagnostic foliaire

Pyriculariose Foliaire										
	Andrano				Ivory					
	2009-2010		2011-2012		2010-2011		2011-2012		2012-2013	
	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation
SiO2	Pas d'effet système	Pas d'effet fertilisation		Pas d'effet fertilisation		*			Pas d'effet système	
N						*				
P						*				
K							*			
Ca							*			
Mg							*			
Na							*			
Cu							*			
Mn							*			
Zn							*			
B							*			
S										*
NNH4										*
NNO3										*
NAlphaAmi								*		

Tableau 34 Récapitulatif covariables de diagnostic foliaire pyriculariose foliaire

Pour la pyriculariose foliaire on peut voir sur le tableau ci-dessus que pour le site Andranomanelatra en 2011-2012 seul l'effet système de culture avait été trouvé significatif et aucune des différentes covariables n'a permis d'expliquer cet effet.

Concernant le site Ivory en 2010-2011 nous nous retrouvons dans le même cas que pour les covariables de prélèvement : c'est-à-dire que toutes les covariables supplantent au moins un des deux effets.

Au contraire en 2011-2012 aucune des covariables n'a permis de supplanter l'effet système de culture ou l'effet fertilisation.

En 2012-2013 seul l'effet fertilisation est significatif, et il a pu être supplanté par la covariable Soufre (S).

Pyriculariose Paniculaire													
	Andrano				Ivory								
	2009-2010		2011-2012		2010-2011		2011-2012		2012-2013				
	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation	Système	Fertilisation			
SiO2		Pas d'effet fertilisation		Pas d'effet fertilisation		*		Pas d'effet système		Pas d'effet fertilisation			
N													
P												*	
K													
Ca												*	
Mg													
Na													
Cu													*
Mn													
Zn												*	
B													
S													
NNH4													
NNO3												*	*
NAlphaAmi													

Tableau 35 Récapitulatif covariables de diagnostic foliaire pyriculariose paniculaire

Pour la pyriculariose paniculaire, le tableau ci-dessus nous montre pour le site Andranomanelatra qu'en 2009-2010 seul l'effet système avait été trouvé significatif précédemment; il n'a pu être supplanté par aucune des différentes covariables.

Il en a été de même pour l'année 2011-2012.

Concernant le site Ivory, en 2010-2011 seul l'effet fertilisation est significatif et il a pu être expliqué par la covariable SiO2.

En 2012-2013 les deux effets ont été trouvés significatifs. Les covariables Potassium (P), Calcium (Ca), Zinc (Zn) et NNO3 ont permis de supplanter l'effet système de culture, alors que les covariables Cuivre (Cu) et NNO3 ont permis de supplanter l'effet fertilisation.

Nous nous attendions à trouver une ou plusieurs covariables qui auraient expliqué les mêmes effets pour les différentes années et les différents sites, mais cela n'a pas été le cas. Peut-être y'a-t-il d'autres covariables à prendre en considération pour refaire les analyses.

5.3.3.3 ANALYSE GENERALE

Afin de visualiser le comportement global de toutes les parcelles et de voir s'il correspond aux conclusions que nous avons pu tirer précédemment, nous réalisons l'étude cette fois-ci sur l'ensemble des parcelles, quelles que soient les années.

En ce qui concerne la fonction de variance, nous avons décidé de la rendre variable pour chaque essai, en effet une régression du log de la variance sur le log de la moyenne sera effectuée pour chaque effet et comme précédemment nous obtiendrons une fonction de variance de même type, mais avec des paramètres différents.

S'agissant de mesures répétées sur les mêmes parcelles, il faudrait à la fois estimer les effets et les corrélations entre erreurs. C'est ici difficile car de la pyriculariose foliaire ou paniculaire nous n'avons au maximum que 2 mesures successives sur une même parcelle. Or ces mesures sont faites sur une même parcelle, mais pas sur les mêmes poquets, car la diagonale qu'ils occupent est changée à chaque nouvelle mesure. Par conséquent, si la variabilité des mesures est essentiellement due à celle des poquets, les mesures peuvent être très peu corrélées, et en pratique considérées comme indépendantes.

Nous vérifions cette hypothèse en comparant la fonction de variance inter-poquets, estimée au cours des analyses préliminaires, à la fonction de variance inter-parcelles, dont seul le coefficient phi est réestimé avec les autres coefficients du modèle linéaire généralisé.

La variance inter-poquets est modélisée par une fonction à deux coefficients k et p :

$$V_1(\mu) = k \mu^p (1 - \mu)^p$$

La variance inter-parcelles (conditionnelle au prédicteur linéaire) est modélisée par une fonction de même forme :

$$V_2(\mu) = \phi \mu^b (1 - \mu)^b$$

Si la variabilité des mesures est essentiellement due à celle des poquets, comme la moyenne par parcelle est obtenue à partir de 10 poquets, nous devons avoir

$$V_2(\mu) = V_1(\mu)/10,$$

soit $b=p$ et $\phi=k/10$.

Nous avons vu précédemment que le coefficient p calculé sur la fonction de variance inter-poquets et adopté comme coefficient b n'était pas remis en cause par la vérification visuelle des résidus. Nous vérifions maintenant que ϕ n'est pas plus grand que $k/10$ en comparant les valeurs estimées pour l'un et l'autre de ces paramètres. k est l'ordonnée à l'origine de la régression linéaire de l'analyse préliminaire, tandis que ϕ est estimé par la procédure Glimmix par la surdispersion (Pearson Chi-Square / DF)

Site	Année	ϕ	$\log(k)$	$k/10$
Andranomanelatra	2009-2010	0,05	-0,015873	0,0984252
	2011-2012	0,05	-0,286831	0,0750639
Ivory	2010-2011	0,06	-0,70945	0,0491915
	2011-2012	0,03	-0,479086	0,0619349
	2012-2013	0,03	-0,441565	0,0643029

Tableau 36 : Comparaison entre les surdispersions observée ϕ et attendue $k/10$

La fonction de variance intra-parcelle divisée par le nombre d'individus (de poquets) est donc supérieure numériquement à la fonction de variance inter-parcelles. Tout se passe comme s'il n'y avait pas d'effet parcelle, donc pas de possibilité de corrélation entre mesures successives. Nous analysons donc l'ensemble comme s'il s'agissait d'essais indépendants, bien que nous ayons des mesures répétées sur les mêmes parcelles.

ANDRANOMANELATRA

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc(Annee)	6	67	2.88	0.0148
Variete	1	67	57.38	<.0001
Systeme	1	67	10.08	0.0023
Systeme*Variete	1	67	3.27	0.0748
Fertilisation	1	67	7.95	0.0063
Fertilisatio*Variete	1	67	0.25	0.6173
Fertilisatio*Systeme	1	67	0.54	0.4639
Fertil*System*Variet	1	67	3.08	0.0839
DateNotationP(Annee)	1	67	37.99	<.0001
Variet*DateNo(Annee)	2	67	5.76	0.0049
System*DateNo(Annee)	2	67	0.07	0.9285
Syst*Vari*Date(Anne)	2	67	1.20	0.3070
Fertil*DateNo(Annee)	2	67	1.68	0.1945
Fert*Vari*Date(Anne)	2	67	0.48	0.6184
Fert*Syst*Date(Anne)	2	67	0.29	0.7525
Fer*Sys*Var*Dat(Ann)	2	67	0.03	0.9698

Le tableau « Type III Tests of Fixed Effects » nous renseigne sur le fait qu'une seule interaction est significative : Variete*DateNotation(Annee), elle regroupe l'interaction Variete année et variete x date. Il ressort également que l'essai, la fertilisation, le système de culture, la variété et enfin le bloc sont tous significatifs, les interactions entre ces facteurs ou entre ces facteurs et la date d'observation ne sont pas significatives

Tableau 37 : Tests Effet Type III

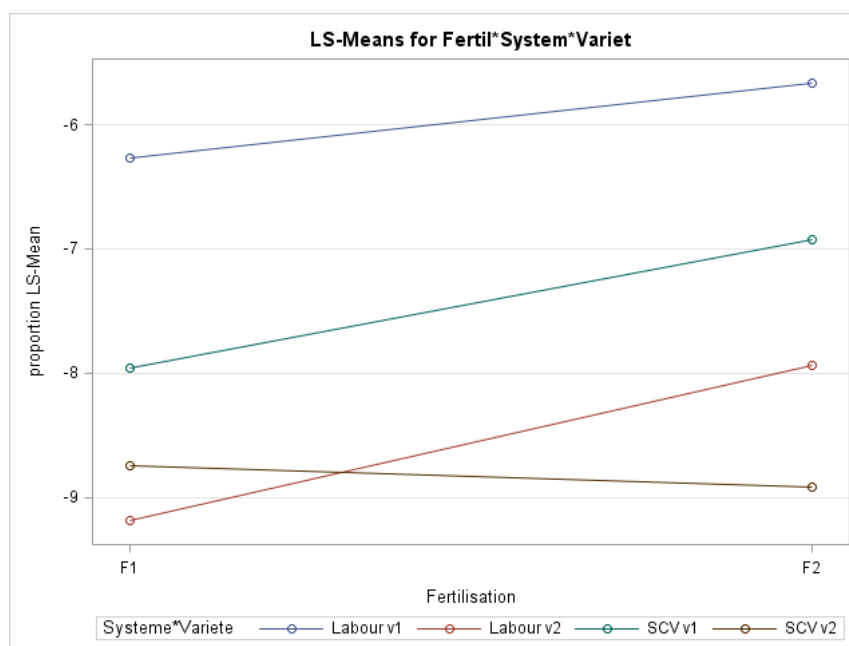


Figure 18 : Evolution du logit de la proportion suivant la variété, la fertilisation et le système de culture

Le graphique ci-dessus nous montre que la variété 1 est plus touchée par la maladie que la variété 2, mais aussi que le système de culture en labour est plus touché que le SCV excepté en Fertilisation 1 pour la variété F2 mais l'interaction entre ces deux facteurs n'est pas ressortie significative.

Enfin plus la fertilisation est azotée plus la maladie est forte, sauf dans le cas de la variété 2 cultivée en SCV ou le taux de pyriculariose diminue légèrement en Fertilisation 2 par rapport à la fertilisation 1. Rappelons que nous sommes à des taux de pyriculariose très faibles, les niveaux des tests asymptotiques sont donc approximatifs.

IVORY

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc(Annee)	9	105	15.02	<.0001
Variete	1	105	18.28	<.0001
Systeme	1	105	60.17	<.0001
Systeme*Variete	1	105	0.96	0.3288
Fertilisation	1	105	42.47	<.0001
Fertilisatio*Variete	1	105	2.36	0.1275
Fertilisatio*Systeme	1	105	2.62	0.1084
Fertil*System*Variet	1	105	0.05	0.8259
DateNotationP(Annee)	2	105	20.42	<.0001
Variet*DateNo(Annee)	4	105	20.67	<.0001
System*DateNo(Annee)	4	105	2.93	0.0241
Syst*Vari*Date(Anne)	4	105	1.38	0.2457
Fertil*DateNo(Annee)	4	105	1.54	0.1964
Fert*Vari*Date(Anne)	4	105	0.39	0.8150
Fert*Syst*Date(Anne)	4	105	2.52	0.0456
Fer*Sys*Var*Dat(Ann)	4	105	1.47	0.2160

Le tableau « Type III Tests of Fixed Effects » nous renseigne sur le fait que plusieurs interactions sont significatives : dans un premier temps on remarque l'interaction triple fertilisation*système*Essai, puis les interactions doubles avec système*Essai, variété*Essai. Ensuite viennent les facteurs Bloc, Variété, Système, Fertilisation et Essai qui sont tous hautement significatifs (p-value <0.0001). La puissance des tests est améliorée par une infection plus élevée par la pyriculariose.

Tableau 38: Test Effet Type III

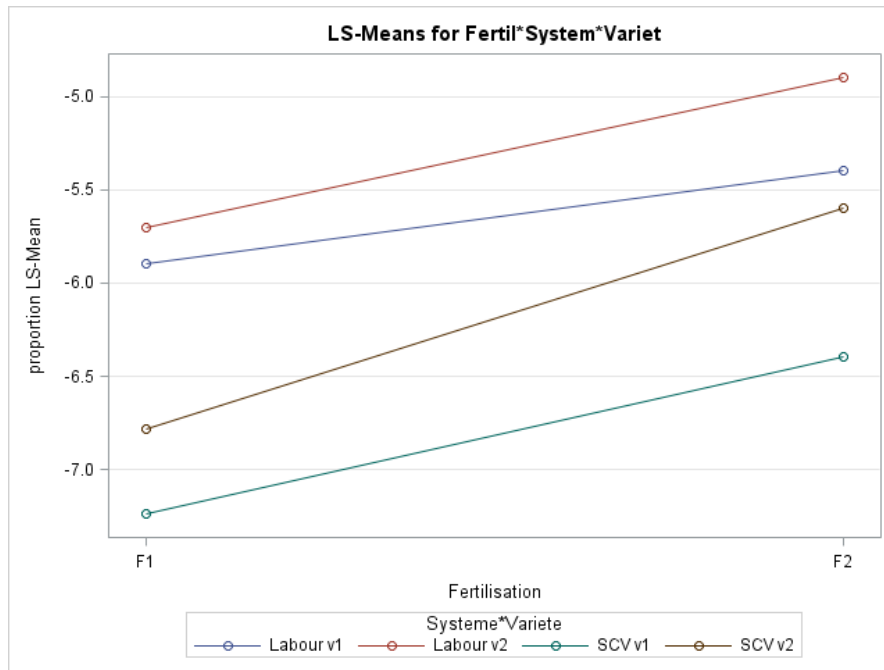


Figure 19 : Evolution du logit de la proportion suivant la variété la fertilisation et le système de culture

Le graphique ci-dessus nous montre que la variété 2 est plus touchée par la maladie que la variété 1, mais aussi que le système de culture en labour est plus touché que le SCV. Enfin plus la fertilisation est azotée est élevée, plus la maladie est forte.

6. CONCLUSION

Ces 4 mois de stage passés au Centre de coopération internationale en recherche agronomique pour le développement de Montpellier ont été très enrichissants.

D'un côté, le travail effectué sur la pyriculariose a permis de mettre en évidence l'absence d'interaction entre les facteurs, systèmes de culture et fertilisation, cette interaction faisait partie des hypothèses de bases du projet GARP ; ce qui signifie que l'hypothèse de base n'a pas été vérifiée.

D'un autre côté le lien azote et pyriculariose a pu être prouvé, en effet l'azote favorise la maladie. Une des autres conclusions tirées de ces études est que le système de culture en labour est plus favorable à l'évolution de la pyriculariose. . Toutefois ce rapport serait dû à une densité de végétation généralement plus forte après labour.

Aucune des covariables que nous avons à notre disposition n'a permis d'expliquer l'effet des facteurs système de culture et fertilisation, peut-être faut-il envisager de s'intéresser à d'autres covariables.

La question qui peut se poser est de savoir si les mêmes résultats auraient pu être obtenus avec une simple analyse de variance sur les logarithmes.

D'un autre côté, ce stage a aussi été enrichissant à un niveau plus personnel, effectivement il m'a permis de mettre en pratique les connaissances acquises tout au long de mon cursus universitaire. En particulier en statistiques avec l'apprentissage de nouvelles méthodes et l'amélioration de mes connaissances du logiciel SAS ; mais aussi en informatique, j'ai exercé mes connaissances du langage SQL et des bases de données.

Enfin l'insertion dans le monde professionnel, en particulier dans un centre de recherche sera bénéfique pour ma future carrière. J'ai également pu voir comment fonctionnaient les interactions entre les personnes travaillant sur le terrain et les statisticiens ; ce qui fut une expérience très positive.

7. ANNEXES

7.1 REQUETES SQL

7.1.1 DONNEES D'INFESTATION DE LA PYRICULARIOSE SUR LES PANICULES

```
SELECT PyriPanículaire.IDParc, PyriPanículaire.Numparc, PyriPanículaire.DateNotationPP, PyriPanículaire.Poquet,
PyriPanículaire.nbPanicles, PyriPanículaire.nbPaniclesInfect,
Round(IIf(IsNull(((P1+P2+P3+P4+P5)/5*nbPaniclesInfect/nbPanicles)),0,((P1+P2+P3+P4+P5)/5*nbPaniclesInfect/nbPanicles)),2) AS SeveritePoquetPonderee,
Round((nbPaniclesInfect/nbPanicles*100),2) AS IncidencePourcentage
FROM PyriPanículaire;
```

```
SELECT SPP.IDParc, SPP.Numparc, SPP.DateNotationPP, SPP.Poquet, P.Bloc, P.Systeme, P.Fertilisation,
P.Variete, SPP.nbPanicles, SPP.nbPaniclesInfect, SPP.SeveritePoquetPonderee, SPP.IncidencePourcentage
FROM Sev_Poquet_Panículaire AS SPP, Parcelle AS P
WHERE SPP.IDParc = P.IDParc;
```

```
SELECT Sev_Poquet_Panículaire.IDParc, Sev_Poquet_Panículaire.Numparc, Sev_Poquet_Panículaire.DateNotationPP,
Avg(Sev_Poquet_Panículaire.nbPanicles) AS NbPaniclesMoyen, Avg(Sev_Poquet_Panículaire.nbPaniclesInfect) AS NbPaniclesInfectMoyen,
Avg(Sev_Poquet_Panículaire.SeveritePoquetPonderee) AS SevPoqMoyen,
Round((Avg(Sev_Poquet_Panículaire.nbPaniclesInfect)/Avg(Sev_Poquet_Panículaire.nbPanicles)*100),2) AS IncidenceParcelleMoyenne,
Round(log(AVG(SeveritePoquetPonderee)),5) AS logMeanSev, Round(log(var(SeveritePoquetPonderee)),5) AS logVarSev
FROM Sev_Poquet_Panículaire
GROUP BY Sev_Poquet_Panículaire.IDParc, Sev_Poquet_Panículaire.Numparc, Sev_Poquet_Panículaire.DateNotationPP;
```

```
SELECT SPF.IDParc, SPF.NumParc, P.Bloc, P.Systeme, P.Fertilisation, P.Variete, P.NumSite, SPF.DateNotationPP,
SPF.IncidenceParcelleMoyenne AS Incidence, SPF.SevPoqMoyen AS Severite, IIf(SevPoqMoyen=0,0,Round(logMeanSev,5)) AS logMeanSev2,
IIf(SevPoqMoyen=0,0,Round(logVarSev,5)) AS logVarSev2
FROM Sev_Parcelle_Panículaire AS SPF, Parcelle AS P
WHERE NbPaniclesInfectMoyen is not null
AND P.IDPARC= SPF.IDPARC
ORDER BY SPF.DateNotationPP, SPF.NumParc;
```

7.1.2 DONNEES D'INFESTATION DE LA PYRICULARIOSE SUR LES FEUILLES

```
SELECT Pyrifoliaire.idpf, Pyrifoliaire.idparc, Pyrifoliaire.numparc, Pyrifoliaire.numpoquet, Pyrifoliaire.datenotationpff,
Pyrifoliaire.nbt AS NbTalles, Pyrifoliaire.nbti AS NbTallesInf,
IIf(IsNull(fn),Null,IIf(IsNull([fn-1]),fn,IIf(IsNull([fn-2]),([fn+[fn-2])/2,IIf(IsNull([fn-3]),([fn+[fn-1]+[fn-2])/3,([fn+[fn-1]+[fn-2]+[fn-3])/4])))) AS Severitel234
FROM Pyrifoliaire;
```

```
SELECT Sev_Talle_Foliaire.IDParc, Sev_Talle_Foliaire.NumParc, Parcelle.Variete, Parcelle.Systeme, Parcelle.Fertilisation,
Parcelle.Annee, Parcelle.NumSite, Sev_Talle_Foliaire.DateNotationPF, Sev_Talle_Foliaire.NumPoquet,
Avg(Sev_Talle_Foliaire.NbTalles) AS NbTallesPoquet, Avg(Sev_Talle_Foliaire.NbTallesInf) AS NbTallesInfPoquet,
Round((Avg(Sev_Talle_Foliaire.NbTallesInf)/Avg(Sev_Talle_Foliaire.NbTalles)*100),2) AS IncidencePourcentage,
Round(iif(IsNull(Avg(Sev_Talle_Foliaire.Severitel234)),0,Avg(Sev_Talle_Foliaire.Severitel234)),2) AS Severitel234Poquet,
Round(iif(IsNull(Avg(Sev_Talle_Foliaire.Severitel234*NbTallesInf/NbTalles)),0,Avg(Sev_Talle_Foliaire.Severitel234*NbTallesInf/NbTalles)),2) AS MoyenneDeSeveritel234Ponderee
FROM Sev_Talle_Foliaire, Parcelle
WHERE Parcelle.IDParc = Sev_Talle_Foliaire.IDParc
GROUP BY Sev_Talle_Foliaire.IDParc, Sev_Talle_Foliaire.NumParc, Parcelle.Variete, Parcelle.Systeme, Parcelle.Fertilisation, Parcelle.Annee,
Parcelle.NumSite, Sev_Talle_Foliaire.DateNotationPF, Sev_Talle_Foliaire.NumPoquet;
```

```
SELECT Parcelle.Variete, Parcelle.Systeme, Parcelle.Fertilisation, Parcelle.Annee, Parcelle.NumSite, Sev_Poquet_Foliaire.DateNotationPF,
Sev_Poquet_Foliaire.IDParc, Sev_Poquet_Foliaire.NumParc, RVG(IncidencePourcentage) AS IncidenceMoyennePourcent, RVG(Severitel234Poquet) AS Severitel234Parcelle,
RVG(MoyenneDeSeveritel234Ponderee) AS Severitel234Ponderee, round(var(MoyenneDeSeveritel234Ponderee),5) AS varSevPond,
IIf((AVG(Severitel234Poquet))=0,0,Round(log(avg(MoyenneDeSeveritel234Ponderee)),4)) AS logMeanSev,
IIf((AVG(Severitel234Poquet))=0,0,Round(log(var(MoyenneDeSeveritel234Ponderee)),4)) AS logVarSev
FROM Sev_Poquet_Foliaire, Parcelle
WHERE Parcelle.IDParc = Sev_Poquet_Foliaire.IDParc
GROUP BY Parcelle.Variete, Parcelle.Systeme, Parcelle.Fertilisation, Parcelle.Annee, Parcelle.NumSite, Sev_Poquet_Foliaire.DateNotationPF,
Sev_Poquet_Foliaire.IDParc, Sev_Poquet_Foliaire.NumParc;
```

```
SELECT SPF.IDParc, SPF.NumParc, P.Bloc, SPF.Systeme, SPF.Fertilisation, SPF.Variete, SPF.Annee, SPF.NumSite, SPF.DateNotationPF,
SPF.IncidenceMoyennePourcent AS Incidence, SPF.Severitel234Parcelle AS Severite, SPF.Severitel234Ponderee AS SeveritePond, logMeanSev, logVarSev
FROM Sev_Parcelle_Foliaire SPF, Parcelle P
WHERE P.IDPARC= SPF.IDPARC
AND IncidenceMoyennePourcent <>0
ORDER BY SPF.Annee, SPF.DateNotationPF, SPF.NumParc;
```

7.2 MACROS SAS

7.2.1 MACRO 1 : IMPORTATION DES DONNEES

```
%macro donneesFoliaire;

/*lieu: Ivory ou Andrano
date: au format 01JAN2013*/

PROC IMPORT OUT= Pyrifol.OBSParcelle
  DATAFILE= "D:\Users\stage102\Documents\Emilie Doge\SAS\Sev_Parcelle_Foliaire2.xls"
  DBMS=EXCELCS REPLACE;
  SCANTEXT=YES;
  USEDATE=YES;
  SCANTIME=YES;
RUN;
quit;

DATA Pyrifol.obsparcelle;
SET Pyrifol.obsparcelle (obs=444);
RUN;

PROC SQL;
  create table Pyrifol.obsparcelle2 as
  select IDPArc, NumParc, BLoc, Systeme,variete, Fertilisation, Annee, NumSite, DateNotationPF, Incidence,
Severite,SeveritePond, logMeanSev, logVarSev, SeveritePond/100 as proportion
  from Pyrifol.obsparcelle
  order by IDPArc,NumParc,bloc,systeme,fertilisation,variete;
quit;

%if &lieu=Andrano %then %do;
  PROC SQL;
    create table Pyrifol.obsparcelle3 as
    select *
    from Pyrifol.obsparcelle2
    where Numsite = 1
    and DateNotationPF = "&date."d;
  ;
  quit;
%end;

%else %if &lieu=Ivory %then
  %do;
  PROC SQL;
    create table Pyrifol.obsparcelle3 as
    select *
    from Pyrifol.obsparcelle2
```

```

        where Numsite = 2
        and DateNotationPF = "&date."d ;
; quit;
%end;

%mend donneesFoliaire;

```

7.2.2 MACRO 2 : PRELIMINAIRES

```

%macro preliminaire;
ods pdf close;
ods pdf file="Analyse préliminaire pyriculariose foliaire &lieu. &date..pdf";
/*****/
/** MISE EN EVIDENCE D'UNE RELATION ENTRE MOYENNE ET VARIANCE          **/
/** JUSTIFIANT LE RECOURS A UN MODELE LINEAIRE GENERALISE              **/
/** ou au moins a une transformation de variable                       **/
/* il s'agit ici de la variance intra-parcelle                          */
/*****/
ods html;
title1 "Relation entre log(moyenne) et log(variance intra-parcelle) de la sévérité moyenne par poquet";
title2 "la pente de la régression est à reporter dans la fonction de variance";
title3 "&lieu. &date.";

symbol1 c=blue v=dot h=0.1 cm i=rl;
proc gplot data=PyriFol.obsparcelle3;
  plot logVarSev*logMeanSev/regeqn;
run;
quit;
ods html close;

/* Transférer les données dans un fichier texte pour les rendre utilisables par R */
PROC EXPORT DATA= PyriFol.obsparcelle3
  OUTFILE= "obsparcelle3 &lieu. &date..txt"
  DBMS=TAB REPLACE;
  PUTNAMES=YES;
RUN;
/*****/
/** VERIFICATION DE LA FONCTION DE VARIANCE D APRES LES GRAPHIQUES CROISANT    ***/
/** LES RESIDUS AVEC LES VALEURS PREDITES (forme en trompette ou pas)          ***/
/* ET D'APRES LA REGRESSION DE LA VALEUR ABSOLUE DES RESIDUS                  ***/
/** SUR LES VALEURS PREDITES AVEC TEST SUR LA PENTE (TEST DE GLEISJER)        ***/
/*****/

title "Tableau des sévérités moyennes par parcelle relevées à &lieu. le &date.";
title2 "ainsi que des proportions des feuilles des différents niveaux 1=terminale 2=subterminale etc.";

```

```

proc tabulate data=PyriFol.obsparcelle3;
  class bloc systeme fertilisation variete NumParc;
  var SeveritePond ;
  tables (bloc all),systeme*fertilisation*variete*(SeveritePond*mean*f=8.4);
run;

title1 "modèle linéaire généralisé avec quasi-vraisemblance fonde sur la relation moyenne-variance de la loi
binomiale";
title2 "avec un effet rang moyen de la feuille à partir de la feuille terminale";
title3 "&lieu. &date.";

proc glimmix data=PyriFol.obsparcelle3 PLOTS=(controlplot pearsonpanel);
  class bloc systeme fertilisation variete NumParc;
  _variance_ = _mu_ * (1-_mu_);
  model proportion = bloc systeme|fertilisation|variete|Essai / link=logit htype=3;
  random _residual_ ;/* équivalent de pscale */
  output out=PyriFol.residus_glim predicted=xbeta pearson=reschi student=STDRESCHI;
run;
quit;

data PyriFol.residus_glim;
  set PyriFol.residus_glim;
  absSTDRESchi=abs(STDRESchi);
run;

title4 "Régression de la valeur absolue du résidu de Pearson sur la valeur prédite";
proc reg data=PyriFol.residus_glim;
  var absstdreschi xbeta;
  model absstdreschi=xbeta;
  ods rtf exclude none;
  plot absstdreschi*xbeta/nostat;
run;
quit;

title1 "modèle linéaire généralisé avec quasi-vraisemblance fondée sur la relation variance = [moyenne*(1-
moyenne)]**1.5";
title2 "avec un effet rang moyen de la feuille à partir de la feuille terminale";
title3 "&lieu. &date.";

proc glimmix data=PyriFol.obsparcelle3 PLOTS=(controlplot pearsonpanel);
  class bloc systeme fertilisation variete NumParc;
  _variance_ = _mu_ **1.5 * (1-_mu_)**1.5;
  model proportion = bloc systeme|fertilisation|variete|Essai / link=logit htype=3;
  random _residual_ ;/* équivalent de pscale */
  output out=PyriFol.residus_glim15 predicted=xbeta pearson=reschi student=STDRESCHI;
run;
quit;

```

```

data PyriFol.residus_glim15;
  set PyriFol.residus_glim15;
  absSTDRESchi=abs(STDRESchi);
run;

title4 "Régression de la valeur absolue du résidu de Pearson sur la valeur prédite";
proc reg data=PyriFol.residus_glim15;
  var absstdreschi xbeta;
  model absstdreschi=xbeta;
  ods rtf exclude none;
  plot absstdreschi*xbeta/nostat;
run;
quit;

title1 "modèle linéaire généralisé avec quasi-vraisemblance fondée sur la relation variance = moyenne2*(1-
moyenne)2";
title2 "avec un effet rang moyen de la feuille à partir de la feuille terminale";
title3 "&lieu. &date.";

/* plus simple avec glimmix */
proc glimmix data=PyriFol.obsparcelle3 PLOTS=(controlplot pearsonpanel);
  class bloc systeme fertilisation variete NumParc;
  _variance_ = _mu_**2 * (1-_mu_)**2;
  model proportion = bloc systeme | fertilisation | variete | Essai / link=logit htype=3;
  random _residual_ ; /* équivalent de pscale */
  output out=PyriFol.residus_glim2 predicted=xbeta pearson=reschi student=STDRESCHI;
run;
quit;

data PyriFol.residus_glim2;
  set PyriFol.residus_glim2;
  absSTDRESchi=abs(STDRESchi);
run;

title4 "Régression de la valeur absolue du résidu de Pearson sur la valeur prédite";
proc reg data=PyriFol.residus_glim2;
  var absstdreschi xbeta;
  model absstdreschi=xbeta;
  ods rtf exclude none;
  plot absstdreschi*xbeta/nostat;
run;
quit;

title4 " ";
ods pdf close;
%mend preliminaire;

```

7.2.3 MACRO 3 : ANALYSE

```
%macro modelefeuille(covar);
ods rtf close;
ods rtf file="Analyse pyriculariose foliaire &lieu. &date. &covar..rtf";

proc tabulate data=PyriFol.obsparcelle3;
    var &covar. proportion;
    class bloc fertilisation systeme variete ;
    tables bloc,fertilisation*systeme*variete*proportion;
run;

title1 "modèle linéaire généralisé avec quasi-vraisemblance fondée sur la relation variance = [moyenne*(1-
moyenne)]**&b.";
title2 "&lieu. &date. &covar.";
title3 "Décomposition de l'effet fertilisation";

%put &b &covar;

proc glimmix data=PyriFol.obsparcelle3 PLOTS=(controlplot pearsonpanel);
    class bloc fertilisation systeme variete ;
    _variance_ = _mu_ **&b. * (1-_mu_) **&b.;
    model proportion = &covar. bloc systeme | fertilisation | variete | Essai / link=logit htype=3;
    random _residual_;
    lsmeans systeme*fertilisation*variete/slice=variete*fertilisation plot=meanplot(sliceby=variete*systeme
join) pdiff;
    output out=PyriFol.residus_glim_m0 predicted=xbeta pearson=reschi student=STDRESCHI;
run;
quit;

ods html close;
ods rtf close;
%mend modelefeuille;
```

7.2.4 MACRO 4 : APPEL DES DIFFERENTES MACROS

```
libname Pyrifol '.';

%let lieu=Andrano;
%let date=08FEB2012;

%include "D:\Users\stage102\Documents\Emilie Doge\SAS\Foliaire Importation Donnees 2.sas";
%donneesFoliaire;

data pyrifol.obsparcelle3;
    set pyrifol.obsparcelle2;
run;

%include "D:\Users\stage102\Documents\Emilie Doge\SAS\Foliaire Glimmix.sas";
%include "D:\Users\stage102\Documents\Emilie Doge\SAS\Foliaire Analyse 2.sas";
%preliminaire;

%let b=1.542334;

%modelefeuille();

/*Suppression variété ou fertilization à valeurs nulles*/
%let lieu=Ivory;
%let date=12JAN2010;

%donneesFoliaire;

PROC SQL;
    create table PyriFol.Obsparcelle3 as
    select *
    from PyriFol.Obsparcelle3
    where (variete = 'v1'
    or variete = 'v2')
    AND (fertilisation ='F1'
    Or fertilisation ='F2');
;

%preliminaire;

%let b=1.539965;

%modelefeuille();

proc tabulate data=PyriFol.Obsparcelle3 out=pyriFol.tabulate;
class variete fertilisation systeme;
var severite;
TABLE severite * ( Mean) , systeme*fertilisation*variete;
run;
```

7.3 SORTIES SAS

7.3.2 ANDRANO PYRICULARIOSE FOLIAIRE 2011-2012

ANALYSE

Model Information	
Data Set	PYRIFOL.OBSPARCELLE3
Response Variable	proportion
Response Distribution	Unknown
Link Function	Logit
Variance Function	$_{\mu}^{**1.555385} * (1-_{\mu})^{**1.555385}$
Variance Matrix	Diagonal
Estimation Technique	Quasi-Likelihood
Degrees of Freedom Method	Residual

Class Level Information		
Class	Levels	Values
Bloc	4	A B C D
Fertilisation	3	F1 F2 Fv
Systeme	2	Labour SCV
Variete	2	v1 v2

Number of Observations Read	95
Number of Observations Used	95

Dimensions	
Covariance Parameters	1
Columns in X	40
Columns in Z	0
Subjects (Blocks in V)	1
Max Obs per Subject	95

Optimization Information	
Optimization Technique	Newton-Raphson
Fisher Scoring	to convergence
Parameters in Optimization	15

Optimization Information	
Lower Boundaries	0
Upper Boundaries	0
Fixed Effects	Not Profiled

Iteration History					
Iteration	Restarts	Evaluations	Objective Function	Change	Max Gradient
0	0	4	54.048750156	.	31.75417
1	0	3	27.8876194	26.16113076	18.68038
2	0	3	14.00561669	13.88200271	10.71608
3	0	3	6.9706624538	7.03495424	5.621274
4	0	3	3.8720822188	3.09858024	2.538941
5	0	3	2.860924477	1.01115774	0.859664
6	0	3	2.6913456125	0.16957886	0.153559
7	0	3	2.6831586273	0.00818699	0.007874
8	0	3	2.6830270667	0.00013156	0.000306
9	0	3	2.6830236796	0.00000339	0.000081
10	0	3	2.683023571	0.00000011	0.000011
11	0	3	2.6830235628	0.00000001	2.979E-6

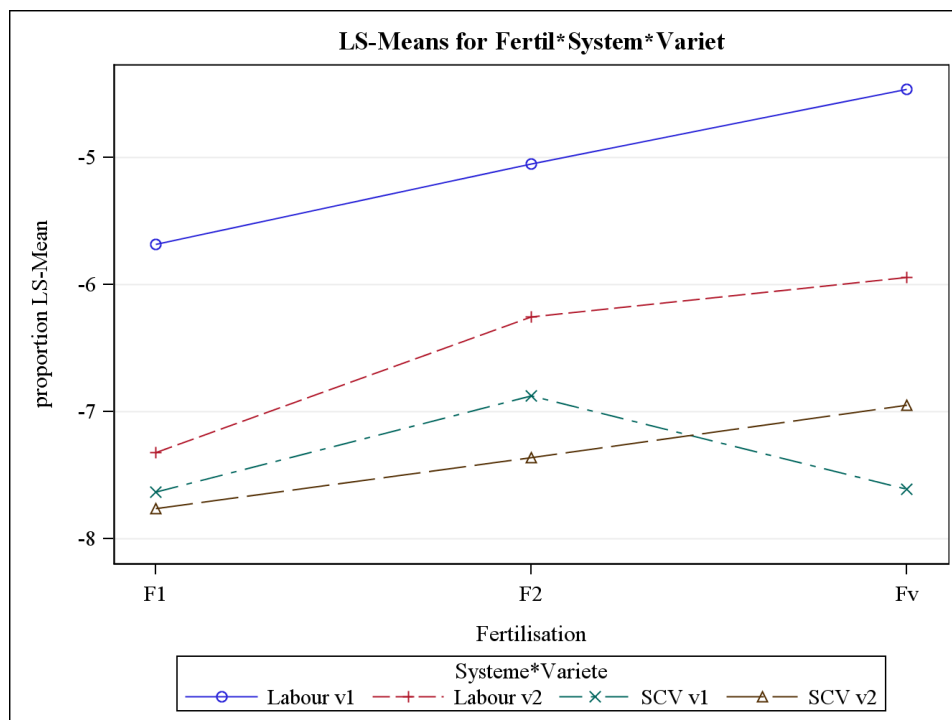
Scoring did not stop.

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics	
-2 Log Quasi-Likelihood	5.37
Quasi-AIC (smaller is better)	35.37
Quasi-AICC (smaller is better)	41.44
Quasi-BIC (smaller is better)	73.67
Quasi-CAIC (smaller is better)	88.67
Quasi-HQIC (smaller is better)	50.85
Pearson Chi-Square	5.58
Pearson Chi-Square / DF	0.07

Type III Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
Bloc	3	80	1.85	0.1451
Systeme	1	80	41.49	<.0001
Fertilisation	2	80	4.28	0.0172
Fertilisatio*Systeme	2	80	1.12	0.3310
Variete	1	80	8.45	0.0047
Systeme*Variete	1	80	8.78	0.0040
Fertilisatio*Variete	2	80	0.40	0.6713
Fertil*System*Variet	2	80	0.79	0.4578

Fertil*System*Variet Least Squares Means							
Fertilisation	Systeme	Variete	Estimate	Standard Error	DF	t Value	Pr > t
F1	Labour	v1	-5.6867	0.3302	80	-17.22	<.0001
F1	Labour	v2	-7.3219	0.4742	80	-15.44	<.0001
F1	SCV	v1	-7.6313	0.5079	80	-15.03	<.0001
F1	SCV	v2	-7.7638	0.5230	80	-14.84	<.0001
F2	Labour	v1	-5.0526	0.2873	80	-17.59	<.0001
F2	Labour	v2	-6.2529	0.3742	80	-16.71	<.0001
F2	SCV	v1	-6.8746	0.4294	80	-16.01	<.0001
F2	SCV	v2	-7.3625	0.4785	80	-15.39	<.0001
Fv	Labour	v1	-4.4660	0.2529	80	-17.66	<.0001
Fv	Labour	v2	-5.9443	0.3496	80	-17.01	<.0001
Fv	SCV	v1	-7.6103	0.5341	80	-14.25	<.0001
Fv	SCV	v2	-6.9493	0.4366	80	-15.92	<.0001



Differences of Fertil*System*Variete Least Squares Means										
Fertilisation	Systeme	Variete	_Fertilisation	_Systeme	_Variete	Estimate	Standard Error	DF	t Value	Pr > t
F1	Labour	v1	F1	Labour	v2	1.6352	0.5775	80	2.83	0.0059
F1	Labour	v1	F1	SCV	v1	1.9445	0.6055	80	3.21	0.0019
F1	Labour	v1	F1	SCV	v2	2.0771	0.6183	80	3.36	0.0012
F1	Labour	v1	F2	Labour	v1	-0.6341	0.4373	80	-1.45	0.1510
F1	Labour	v1	F2	Labour	v2	0.5662	0.4987	80	1.14	0.2596
F1	Labour	v1	F2	SCV	v1	1.1879	0.5414	80	2.19	0.0311
F1	Labour	v1	F2	SCV	v2	1.6757	0.5811	80	2.88	0.0050
F1	Labour	v1	Fv	Labour	v1	-1.2208	0.4155	80	-2.94	0.0043
F1	Labour	v1	Fv	Labour	v2	0.2576	0.4805	80	0.54	0.5934
F1	Labour	v1	Fv	SCV	v1	1.9236	0.6272	80	3.07	0.0029
F1	Labour	v1	Fv	SCV	v2	1.2626	0.5471	80	2.31	0.0236
F1	Labour	v2	F1	SCV	v1	0.3094	0.6946	80	0.45	0.6572
F1	Labour	v2	F1	SCV	v2	0.4419	0.7057	80	0.63	0.5330
F1	Labour	v2	F2	Labour	v1	-2.2693	0.5541	80	-4.10	0.0001
F1	Labour	v2	F2	Labour	v2	-1.0689	0.6038	80	-1.77	0.0805
F1	Labour	v2	F2	SCV	v1	-0.4473	0.6394	80	-0.70	0.4863
F1	Labour	v2	F2	SCV	v2	0.04057	0.6734	80	0.06	0.9521
F1	Labour	v2	Fv	Labour	v1	-2.8559	0.5371	80	-5.32	<.0001
F1	Labour	v2	Fv	Labour	v2	-1.3776	0.5888	80	-2.34	0.0218

Differences of Fertil*System*Variet Least Squares Means										
Fertilisation	Systeme	Variete	_Fertilisation	_Systeme	_Variete	Estimate	Standard Error	DF	t Value	Pr > t
F1	Labour	v2	Fv	SCV	v1	0.2884	0.7135	80	0.40	0.6871
F1	Labour	v2	Fv	SCV	v2	-0.3726	0.6443	80	-0.58	0.5647
F1	SCV	v1	F1	SCV	v2	0.1325	0.7288	80	0.18	0.8561
F1	SCV	v1	F2	Labour	v1	-2.5786	0.5832	80	-4.42	<.0001
F1	SCV	v1	F2	Labour	v2	-1.3783	0.6306	80	-2.19	0.0318
F1	SCV	v1	F2	SCV	v1	-0.7567	0.6648	80	-1.14	0.2585
F1	SCV	v1	F2	SCV	v2	-0.2688	0.6975	80	-0.39	0.7010
F1	SCV	v1	Fv	Labour	v1	-3.1653	0.5670	80	-5.58	<.0001
F1	SCV	v1	Fv	Labour	v2	-1.6869	0.6162	80	-2.74	0.0076
F1	SCV	v1	Fv	SCV	v1	-0.02093	0.7363	80	-0.03	0.9774
F1	SCV	v1	Fv	SCV	v2	-0.6819	0.6695	80	-1.02	0.3114
F1	SCV	v2	F2	Labour	v1	-2.7112	0.5964	80	-4.55	<.0001
F1	SCV	v2	F2	Labour	v2	-1.5108	0.6428	80	-2.35	0.0212
F1	SCV	v2	F2	SCV	v1	-0.8892	0.6765	80	-1.31	0.1924
F1	SCV	v2	F2	SCV	v2	-0.4013	0.7086	80	-0.57	0.5727
F1	SCV	v2	Fv	Labour	v1	-3.2978	0.5806	80	-5.68	<.0001
F1	SCV	v2	Fv	Labour	v2	-1.8195	0.6288	80	-2.89	0.0049
F1	SCV	v2	Fv	SCV	v1	-0.1535	0.7469	80	-0.21	0.8377
F1	SCV	v2	Fv	SCV	v2	-0.8145	0.6810	80	-1.20	0.2353
F2	Labour	v1	F2	Labour	v2	1.2003	0.4714	80	2.55	0.0128
F2	Labour	v1	F2	SCV	v1	1.8220	0.5163	80	3.53	0.0007
F2	Labour	v1	F2	SCV	v2	2.3098	0.5578	80	4.14	<.0001
F2	Labour	v1	Fv	Labour	v1	-0.5867	0.3822	80	-1.53	0.1288
F2	Labour	v1	Fv	Labour	v2	0.8917	0.4521	80	1.97	0.0520
F2	Labour	v1	Fv	SCV	v1	2.5577	0.6057	80	4.22	<.0001
F2	Labour	v1	Fv	SCV	v2	1.8967	0.5223	80	3.63	0.0005
F2	Labour	v2	F2	SCV	v1	0.6216	0.5693	80	1.09	0.2781
F2	Labour	v2	F2	SCV	v2	1.1095	0.6071	80	1.83	0.0714
F2	Labour	v2	Fv	Labour	v1	-1.7870	0.4512	80	-3.96	0.0002
F2	Labour	v2	Fv	Labour	v2	-0.3086	0.5117	80	-0.60	0.5481
F2	Labour	v2	Fv	SCV	v1	1.3574	0.6514	80	2.08	0.0404
F2	Labour	v2	Fv	SCV	v2	0.6964	0.5747	80	1.21	0.2292
F2	SCV	v1	F2	SCV	v2	0.4879	0.6426	80	0.76	0.4500
F2	SCV	v1	Fv	Labour	v1	-2.4086	0.4980	80	-4.84	<.0001
F2	SCV	v1	Fv	Labour	v2	-0.9303	0.5534	80	-1.68	0.0966

Differences of Fertil*System*Variet Least Squares Means										
Fertilisation	Systeme	Variete	_Fertilisation	_Systeme	_Variete	Estimate	Standard Error	DF	t Value	Pr > t
F2	SCV	v1	Fv	SCV	v1	0.7357	0.6846	80	1.07	0.2857
F2	SCV	v1	Fv	SCV	v2	0.07471	0.6121	80	0.12	0.9032
F2	SCV	v2	Fv	Labour	v1	-2.8965	0.5408	80	-5.36	<.0001
F2	SCV	v2	Fv	Labour	v2	-1.4182	0.5922	80	-2.39	0.0190
F2	SCV	v2	Fv	SCV	v1	0.2479	0.7164	80	0.35	0.7303
F2	SCV	v2	Fv	SCV	v2	-0.4132	0.6474	80	-0.64	0.5252
Fv	Labour	v1	Fv	Labour	v2	1.4783	0.4310	80	3.43	0.0010
Fv	Labour	v1	Fv	SCV	v1	3.1444	0.5901	80	5.33	<.0001
Fv	Labour	v1	Fv	SCV	v2	2.4833	0.5042	80	4.93	<.0001
Fv	Labour	v2	Fv	SCV	v1	1.6660	0.6375	80	2.61	0.0107
Fv	Labour	v2	Fv	SCV	v2	1.0050	0.5589	80	1.80	0.0759
Fv	SCV	v1	Fv	SCV	v2	-0.6610	0.6891	80	-0.96	0.3403

Tests of Effect Slices for Fertil*System*Variet Sliced By Fertilisatio*Variete					
Fertilisation	Variete	Num DF	Den DF	F Value	Pr > F
F1	v1	1	80	10.31	0.0019
F1	v2	1	80	0.39	0.5330
F2	v1	1	80	12.45	0.0007
F2	v2	1	80	3.34	0.0714
Fv	v1	1	80	28.39	<.0001
Fv	v2	1	80	3.23	0.0759

