



Institut de Santé Publique, d'Epidémiologie et de Développement
Université de Bordeaux

2^{de} année du Master Santé Publique, Parcours Biostatistique

Promotion 2018-2019

Développement d'une démarche d'analyses statistiques de données d'un réseau pluriannuel d'expérimentations systèmes de culture en vergers de pêchers



BOSTAL Constance

Stage réalisé du 05mars 2019 au 30aout 2019

A l'Institut National de la Recherche Agronomique Provence-Alpes-Côte d'Azur,

Unité Plantes et Systèmes de cultures Horticoles (PSH)

Maîtres de stage : Mohamed Memmah, chercheur

Julie Borg, ingénieur de recherche

Daniel Plénet, chercheur

Francoise Lescourret, directrice de recherche

Tuteur universitaire : Robin Genuer, maître de conférences en statistique

Lecteur : Valérie Kiewsky, professeure agrégée de mathématiques et informatique

Ce rapport de stage a été réalisé dans le cadre du projet ECOVERGER, action pilotée par le ministère de l'Agriculture et de l'alimentation et le ministère de la Transition écologique et solidaire, avec l'appui financier de l'Agence française pour la biodiversité dans le cadre de l'APR "Résistance et pesticides" grâce aux crédits issus de la redevance pour pollutions diffuses attribués au financement du plan Ecophyto.

This work has been carried out as part of the ECOVERGER project. This action is led by the Ministry for Agriculture and Food and the Ministry for an Ecological and Solidary Transition, with the financial support of the French Biodiversity Agency on "Resistance and Pesticides" research call, with the fees for diffuse pollution coming from the Ecophyto plan.



Remerciements

Je tiens tout d'abord à remercier mes maitres de stage, M. Mohamed Memmah, Mme Julie Borg, M. Daniel Plénet et Mme Françoise Lescourret, d'avoir bien voulu m'accueillir au sein de leur équipe. Je les remercie sincèrement pour leurs conseils, leur bienveillance et leurs attentions tout au long de mon stage.

Je remercie également mes enseignants, M. Robin Genuer, M. Boris Hejblum et Mme Cécile Proust-Lima, pour leur disponibilité et leurs conseils dans la réalisation de mes missions.

Je souhaite remercier tous les membres de l'unité PSH et autres membres de l'INRA d'Avignon que j'ai pu côtoyer durant ces six derniers mois, pour leur sympathie. Merci à mes camarades de bureau et autres stagiaires, Amira, Margaux, Mélanie, Dorine, Fanny, Cécile et Nathalie avec qui j'ai partagé cette expérience et de très sympathiques moments. Je leur souhaite de réussir dans leurs études et dans leurs futures carrières.

Merci également aux membres de ma famille pour leur soutien, notamment à mes parents pour leurs encouragements, et mon oncle, ma tante et mes cousines pour m'avoir accueilli dans leur région.

Sommaire

| | |
|---|----|
| Remerciements | 3 |
| Sommaire | 4 |
| Introduction..... | 5 |
| 1. Présentation de la structure d'accueil | 7 |
| 1.1. L'INRA | 7 |
| 1.1.1. Présentation globale..... | 7 |
| 1.1.2. Les centres..... | 7 |
| 1.2. L'INRA PACA | 7 |
| 1.3. L'Unité Plantes et Systèmes de cultures Horticoles | 8 |
| 1.3.1. Rôle et travaux de l'unité | 8 |
| 1.3.2. Les équipes de recherches | 9 |
| 2. Contexte et objectifs | 10 |
| 2.1. Présentation du projet EcoPêche..... | 10 |
| 2.1.1. Mise en place et informations techniques | 10 |
| 2.1.2. Présentation des sites et systèmes | 11 |
| 2.1.3. L'expérimentation en arboriculture | 12 |
| 2.2. Objectif du stage | 12 |
| 3. Matériels et Méthodes..... | 13 |
| 3.1. Présentation des données..... | 13 |
| 3.1.1. Description des bases de données | 13 |
| 3.1.2. Description des données | 13 |
| 3.2. Préparation des données | 14 |
| 3.2.1. Importation et fusion des bases de données | 14 |
| 3.2.2. Choix des sites et de la période..... | 14 |
| 3.2.3. Uniformisation des types de système | 14 |
| 3.2.4. Ajout d'identifiants et variable temporelle | 14 |
| 3.2.5. Recodage des leviers qualitatifs | 15 |
| 3.3. Méthodes d'analyses utilisées | 15 |
| 3.3.1. Analyse descriptive..... | 15 |
| 3.3.2. Analyses factorielles | 15 |
| 5.1.1. Analyse de données longitudinales | 18 |
| 5.2. Gestion des données manquantes..... | 19 |
| 5.3. Aspect programmation..... | 20 |

| | | |
|--------|--|----|
| 5.3.1. | Package et fonction R utilisés | 20 |
| 5.3.2. | Optimisation des graphiques sous R | 20 |
| 6. | Résultats | 20 |
| 6.1. | Analyses descriptives..... | 20 |
| 6.1.1. | Caractéristiques générales des systèmes..... | 20 |
| 6.1.2. | Suivi des systèmes | 21 |
| 6.1.3. | Réduction des intrants | 21 |
| 6.1.4. | Utilisation des leviers | 22 |
| 6.1.5. | Comparaison des performances selon les systèmes..... | 23 |
| 6.2. | Analyses multidimensionnelles | 24 |
| 6.2.1. | Quelques indicateurs de performances et variables agronomiques | 24 |
| 6.2.2. | Leviers qualitatifs..... | 26 |
| 6.2.3. | Leviers et indicateurs de performance..... | 28 |
| 6.3. | Analyses longitudinales ou groupées | 29 |
| 6.3.1. | Analyses longitudinales | 29 |
| 7. | Discussion | 32 |
| 7.1. | Réflexions sur les méthodes..... | 32 |
| 7.1.1. | Autres méthodes d'analyse envisagées | 32 |
| 7.1.2. | Faibles corrélations | 33 |
| 7.1.3. | Modèle mixte non linéaire à long terme..... | 33 |
| 7.1.4. | Débat sur la gestion des données manquantes | 33 |
| 7.1.5. | Choix des individus | 33 |
| 7.2. | Réflexions sur les résultats | 34 |
| | Conclusion | 37 |
| | Conclusion du stage..... | 37 |
| | Conclusion personnelle | 37 |
| | Bibliographie..... | 39 |
| | Annexes | 41 |

Introduction

Initié en 2008, le Plan Écophyto I avait pour objectif de réduire de 50 % l'utilisation des produits phytopharmaceutiques d'ici 2018. Cet objectif n'ayant pas été atteint, ce projet a été remanié et

relancé avec le même objectif d'ici 2025 (Écophyto 2+). Les pratiques agronomiques actuelles, notamment l'usage important de produits phytosanitaires pour certains types de cultures peuvent avoir de grands impacts sur les produits de consommation qui en résultent et sur leurs environnements. Les consommateurs ou personnes résidant à proximité de ces cultures peuvent être confrontés à des résidus de ces traitements, ce qui peut avoir des effets néfastes sur leur santé. La réduction des produits phytosanitaires en agriculture est un réel enjeu dans le cadre de la santé publique, ainsi qu'une priorité pour la sauvegarde de la biodiversité.

Dans le cadre de ma formation, à savoir un Master en Santé Publique spécialité Biostatistique parcours biostatistiques, je devais réaliser un stage de 4 mois minimum. J'ai choisi de réaliser mon stage de 6 mois à l'unité Plantes et Systèmes de cultures Horticoles de l'INRA d'Avignon dont l'offre de stage m'a semblé appropriée à mon parcours universitaire et très intéressante car j'avais la possibilité de mettre en pratique mes acquis et également d'en apprendre davantage sur la recherche agronomique.

L'intitulé du stage était «développement d'une démarche d'analyses statistiques de données d'un réseau pluriannuel d'expérimentations systèmes de culture en vergers de pêchers». L'objectif était de proposer et mettre en œuvre une démarche statistique adaptée afin d'exploiter au mieux les données du projet EcoPêche, projet d'expérimentations de systèmes de culture innovants de pêche et nectarine, implantés sur sept sites de 2013 à 2018 et visant une réduction de 50 % de l'utilisation des pesticides comme préconisé par le plan Écophyto. Dans ce rapport, je présenterai, dans un premier temps, la structure d'accueil. Ensuite, je décrirai plus précisément le contexte et l'objectif de ce stage. Puis, je détaillerai les différentes étapes des missions que j'ai réalisées en présentant les méthodes employées, leurs réalisations et résultats obtenus ainsi que la discussion des méthodes, résultats et perspectives.

1. Présentation de la structure d'accueil

Mon stage s'est déroulé au sein de l'unité Plantes et Systèmes de cultures Horticoles (PSH) situé à l'INRA PACA.

1.1. L'INRA

1.1.1. Présentation globale

L'Institut national de la recherche agronomique, connu sous son acronyme INRA, est un organisme national de recherche scientifique dont les travaux concernent 3 domaines : l'alimentation, l'agriculture et l'environnement.

Fondé en 1946, il est le premier institut de recherche agronomique en Europe, et le deuxième en sciences agricoles dans le monde. Il dispose du statut d'Établissement Public à caractère Scientifique et Technologique (EPST), et est sous la double tutelle du Ministère de la Recherche et du Ministère de l'Agriculture.

Les recherches de l'INRA sont déterminées par l'évolution permanente des questionnements scientifiques ainsi que par les défis planétaires associés à l'agronomie et à l'agriculture tels que le changement climatique, l'épuisement des ressources fossiles et la nutrition humaine.

L'INRA a pour objectif de répondre aux problématiques des défis planétaires actuels et futurs associés à l'agronomie et à l'agriculture dans une perspective de développement durable, tels qu'assurer la nutrition humaine et la valorisation des territoires.

Les principales missions de l'INRA sont de produire et diffuser des connaissances scientifiques, de contribuer à l'élaboration de la stratégie nationale de recherche, à l'innovation, à la formation et par la recherche, éclairer les politiques publiques et les décisions des acteurs privés, participer aux débats sur la Science en société, et promouvoir éthique et déontologie de la recherche.

En 2017, l'INRA employait 7 656 agents titulaires et 2 569 agents contractuels, et ses ressources s'élevaient à 850,89 millions d'euros dont 80,4 % issues des subventions pour charges de service public⁽¹⁾.

Dès 2020, l'INRA fusionnera avec l'institut national de recherche en sciences et technologies pour l'environnement et l'agriculture (IRSTEA), et deviendra désormais l'institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE).

1.1.2. Les centres

A ce jour, l'INRA comporte 17 centres de recherche régionaux¹ et plus de 150 sites en France métropolitaine et Outre-mer. Ceux-ci œuvrent sur une vingtaine de pôles prioritaires, conduits au sein de treize départements scientifiques, s'impliquant à la fois au niveau régional et international. Le siège social est situé à Paris.

1.2. L'INRA PACA

Le centre Inra Provence-Alpes-Côte d'Azur a été créé en 2010 en fusionnant les centres INRA d'Avignon (Vaucluse) et de Sophia-Antipolis (Alpes-Maritimes).

¹ Suite à des fusions, on compte 17 centres régionaux contre 21 en 2008.

Les recherches qui y sont menées sont tournées vers l'agro-écologie¹ des systèmes de culture, qu'ils soient sous serres et en vergers, et la modélisation de l'impact au niveau régional du changement climatique à l'échelle du paysage. Il accorde une attention toute particulière aux enjeux spécifiques liés aux territoires méditerranéens.

En 2017, ce centre a bénéficié de 61,4 millions d'euros de budget dont 9,5 millions en ressources propres. Le centre de Provence-Alpes-Côte d'Azur se place au 4ème rang national parmi les 17 centres INRA. Il mobilise 7,8 % des effectifs de l'INRA, et l'ensemble de ses sites rassemblent 5,9 % des capacités de recherche en agronomie disponibles en France.

Il est à la fois soutenu, d'une part par des partenaires académiques (notamment des universités d'Aix-Marseille, d'Avignon et de Nice-Sophia-Antipolis) et de recherche publique (par exemple, le centre national de la recherche scientifique, l'institut national de la santé et de la recherche médicale, l'institut de recherche pour le développement et le commissariat à l'énergie atomique et aux énergies alternatives) et d'autre part par des partenaires institutionnels et territoriaux (par exemple la région Provence-Alpes-Côte d'Azur, le conseil départemental du Vaucluse et le conseil départemental des Alpes Maritimes)(3).

Il dispose de 2 centres de ressources biologiques, 3 plateformes, 10 plateaux techniques et de 20 000 m² de serres et de chambres de culture dédiées au végétal, ainsi que de 3 domaines comportant au total 180 ha de surfaces expérimentales essentiellement dédiées à l'arboriculture. Il rassemble un millier d'agents répartis dans 26 unités, localisés sur les deux principaux sites (Avignon et Sophia-Antipolis) et dans huit autres sites (Aix-en-Provence, Gotheron, Le Cap d'Antibes, Nice, Les Vignères, Manduel, l'Amarine et Marseille).

L'INRA PACA est structuré en 3 pôles thématiques : le pôle Adaptation au Changement Global (ACG), le pôle Santé des Plantes (PSP) et le pôle Production Horticole Intégrée (PHI).

1.3. L'Unité Plantes et Systèmes de cultures Horticoles

1.3.1. Rôle et travaux de l'unité

Les travaux de recherche entrepris par l'unité PSH s'inscrivent dans le pôle "Production Horticole Intégrée" du centre INRA PACA. Cette unité a pour mission de contribuer à la mise au point de systèmes de culture des fruits et légumes ainsi que de scénarios paysagers en zone méditerranéenne visant à améliorer la qualité des produits et le respect de l'environnement.

Pour ce faire, il est nécessaire de comprendre et quantifier l'impact des pratiques agricoles, de l'organisation du paysage et des facteurs de l'environnement sur le fonctionnement de la plante, tout en prenant en compte l'impact des populations d'auxiliaires² et de bio-agresseurs³ sur les cultures. Cet objectif est atteint en développant des modèles permettant d'analyser la complexité et d'évaluer

¹ L'agro-écologie est une manière de concevoir des systèmes de production en s'appuyant sur les capacités naturelles des écosystèmes, tout en visant à diminuer les pressions exercées sur l'environnement et à préserver les ressources naturelles. Il s'agit d'utiliser au maximum la nature comme facteur de production(2).

² « Animal prédateur ou parasite qui, par son mode de vie, apporte son concours à la destruction de ravageurs nuisibles aux cultures »(4)

³ « Organisme vivant appartenant au règne animal ou végétal, ainsi que les virus, bactéries ou autres agents pathogènes, dont la présence n'est pas souhaitée soit parce qu'il est considéré comme un ennemi des végétaux ou des produits végétaux, soit parce qu'il produit un effet néfaste pour l'homme, les animaux ou l'environnement »(4)

les performances des systèmes étudiés à différentes échelles, afin de concevoir des scénarios techniques et paysagers adaptés à la production intégrée.

Les travaux de l'unité PSH combinent des approches expérimentales et de modélisation, abordant ainsi l'agronomie, l'écophysiologie, la génétique, la physiologie, l'écologie des populations et communautés, l'écologie chimique et l'écotoxicologie à de multiples échelles (des cellules d'une plante jusqu'au paysage).

Les études concernent les plantes horticoles, principalement les cultures de tomates et de laitues en maraîchage, ainsi que les vergers de pêchers et de pommiers en arboriculture.

L'unité possède des installations expérimentales diversifiées dont des vergers, des serres, des chambres de culture, des enceintes climatiques et un insectarium, ainsi qu'un plateau technique d'analyse de la composition des végétaux et une plateforme de biologie moléculaire. Il est donc possible de planifier l'ensemble des recherches sur le site, que ce soient les expérimentations ou les analyses biologiques.

La transformation de l'horticulture conventionnelle vers des systèmes conciliant les performances agronomiques et écologiques nécessite un travail complexe car combinant des connaissances scientifiques et l'utilisation de modèles permettant d'explorer et de concevoir des scénarios techniques innovants pour s'adapter à ces enjeux. Pour pouvoir mener à bien cet objectif, l'unité PSH est structurée en 4 équipes de recherches.

1.3.2. Les équipes de recherches

Dans le contexte actuel, l'équipe 1 « Croissance Architecture et Qualité » (CRAQ) étudie l'impact des pratiques culturales, des facteurs abiotiques¹ et du génotype sur la croissance des plantes horticoles, dans le but d'optimiser l'acquisition des ressources, d'assurer la croissance de la plante et de ses organes dans des conditions non optimales, tout en améliorant la qualité des produits frais ou transformés.

L'équipe 2 « Régulation des Bioagresseurs² par les Plantes » (RBP) a la double vocation de générer de nouvelles connaissances concernant les processus associés aux interactions entre les plantes horticoles et les bioagresseurs dans un ensemble de contextes et de proposer des stratégies de gestion des cultures horticoles qui optimisent la défense des plantes et limitent la gravité des symptômes causés par les bioagresseurs.

Le cadre général des travaux de l'équipe 3 « Contrôle Biologique par Conservation » (CBC) est celui de la lutte biologique par conservation³ en vergers dans un contexte de réduction des pesticides.

Les recherches de l'équipe 4, « Conception et Evaluation de Systèmes de Culture horticoles et d'Idéotypes » (CESCI), ont pour objectif de concevoir de nouveaux modes de production durable et économiquement compétitifs répondant aux enjeux environnementaux actuels. Cette équipe mobilise une démarche systémique pour comprendre les différentes composantes des agro-écosystèmes horticoles en prenant en compte l'action des pratiques culturales et des différentes échelles spatiales

¹ Les facteurs abiotiques concernent la matière inerte, tels que la lumière, la température, l'eau, l'air, le sol (facteurs physiques et chimiques), le vent, le feu. Fait opposition aux facteurs biotiques.

² Le terme bioagresseur regroupe tous les organismes pouvant engendrer des dommages sur les cultures. Il peut s'agir d'agents pathogènes responsables de maladies, de ravageurs ou de plantes adventices(5).

³ La lutte biologique par conservation vise à promouvoir le contrôle des ravageurs par la communauté des auxiliaires naturellement présente dans les agro-écosystèmes, en combinaison avec d'autres pratiques de contrôle biologique.

et temporelles. Cette démarche vise ainsi à valoriser les fonctions écologiques afin de réduire l'usage des intrants de synthèse. C'est au sein de cette équipe que le stage a été effectué.

2. Contexte et objectifs

Les travailleurs du milieu agricole sont les plus exposés aux pesticides et donc plus à risques d'en subir les effets néfastes. Bien que le sujet soit plus complexe à traiter, ces effets peuvent également atteindre les consommateurs des produits traités(6). Les pesticides employés en agriculture laissent des résidus aussi bien dans l'environnement (sols, milieux aquatiques, air) que sur les denrées alimentaires produites. L'impact des pesticides sur la santé est un des principaux sujets de recherche actuel. Afin de préserver la santé des populations ainsi que la biodiversité, la réduction des produits phytopharmaceutiques¹ est une nécessité.

D'une part, l'utilisation des produits phytosanitaires impacte à la fois la santé humaine(6) et la biodiversité(8), et d'autre part, il est nécessaire de produire suffisamment pour couvrir les besoins alimentaires, tout en préservant la santé et l'environnement. Dans ce contexte, le plan Ecophyto a été mis en place en 2008 suite au Grenelle de l'environnement. Ce plan a toujours le même objectif après plusieurs remaniements et une prolongation du délai initialement prévu²: diminuer de 50% le recours aux produits phytosanitaires d'ici à 2025(9).

La pêche arrive en deuxième position des cultures arboricoles les plus traitées (après la pomme), avec un IFT moyen de 20(10). La moitié de l'IFT pêche concerne les fongicides pour contrôler 3 maladies : la cloque, la moniliose et l'oïdium (les 3 maladies principales du pêcher et pour lesquelles peu ou pas de moyens de lutte alternatifs existent)(11). L'utilisation de produits phytosanitaires est alors quasi-systématique. La résistance des bioagresseurs aux produits phytosanitaires employés se développant suite à l'utilisation répétée des traitements(12), les systèmes conventionnels devront modifier leur pratiques.

Afin de rendre possible la diminution des IFT en verger de pêcher, il semble indispensable de penser au-delà de l'itinéraire technique, à l'échelle du système de culture et de l'organisation du verger.

2.1. Présentation du projet EcoPêche

2.1.1. Mise en place et informations techniques

Le projet EcoPêche est un projet d'expérimentations de systèmes de culture innovants de pêche - nectarine visant une réduction de 50% de l'utilisation des pesticides. Il s'agit d'un projet multisite et

¹ Le terme produit phytopharmaceutique est synonyme du terme produit phytosanitaire, dont la définition est : « Substance active ou préparation destinée, selon les cas à protéger des végétaux ou des produits végétaux contre tout organisme nuisible ou à prévenir son action, exercer une action sur les processus vitaux des végétaux, pour autant qu'il ne s'agisse pas de substances nutritives, assurer la conservation des produits végétaux, pour autant que ces substances ou produits ne fassent pas l'objet de dispositions communautaires particulières concernant les agents conservateurs, détruire des végétaux ou des parties de végétaux indésirables, prévenir une croissance indésirable des végétaux, à l'exception des algues à moins que les produits ne soient appliqués sur le sol ou l'eau pour protéger des végétaux »(7)

² Cet objectif devait initialement être atteint en 2018. A cette date, une légère baisse avait été constatée, mais les 50% de réduction préconisée étaient loin d'être atteints.

pluriannuel. Le but de ce projet est de trouver un compromis entre réduction des pesticides et maintien d'une marge économique acceptable pour les producteurs.

Des vergers de pêcheurs ont été implantés sur sept sites d'essais systèmes en région méditerranéenne, plus précisément dans la Drome, le Gard, le Vaucluse et les Pyrénées-Orientales (cartographie disponible en Annexe 2). En terme de production nationale, l'Occitanie est la première région (47%) de la production suivie par la région Provence-Alpes-Côte-D'azur (32 %) et la région Auvergne-Rhône-Alpes (16 %)(chiffres de 2017)(15). Sur la période du projet, c'est-à-dire 2013 à fin 2018 (le projet sera par la suite prolongé), de nombreuses données qualitatives et quantitatives ont été enregistrées chaque année dans les différents sites pour caractériser les performances agronomiques, techniques et économiques, ainsi que l'état sanitaire des vergers(13).

2.1.2. Présentation des sites et systèmes

Chaque site dispose d'un système de culture conventionnel (système de référence), ainsi que d'un ou deux systèmes économes en produits phytosanitaires, et parfois d'un système conduit en agriculture biologique. Dans ce rapport, les abréviations REF, ECO et BIO seront couramment utilisées pour définir les types de systèmes. En totalité, on dispose de 3 systèmes BIO, de 7 systèmes REF et de 9 systèmes ECO.

| Sites | Référence (REF) | Econome en intrants (ECO) | Agriculture biologique (BIO) |
|---------------|-----------------|---------------------------|------------------------------|
| INRA Avignon | 1 | 2 | 0 |
| INRA Gotheron | 1 | 1 | 1 |
| CTIFL | 1 | 1 | 1 |
| SEFRA | 1 | 1 | 1 |
| SERFEL 1 | 1 | 1 | 0 |
| SERFEL 2 | 1 | 1 | 0 |
| Sica Centrex | 1 | 2 | 0 |

Tableau 1 : Nombre de système expérimentaux selon les sites et le type de système

Il est à noter que les pratiques agricoles, même en système conventionnel, diffèrent d'un site à l'autre car les règles de conduite en culture de pêcheurs sont assez flexibles, aussi bien en termes de choix de leviers¹ qu'en quantités d'intrants employées (irrigation, fertilisation, indicateur de fréquence de traitements phytosanitaires² (IFT)), et sont souvent adaptées par les arboriculteurs selon les différents cas. Les sites d'expérimentations sont très différents de par leur localisation géographique et leurs conditions pédoclimatiques. En conséquence, les sites n'ont pas les mêmes variétés ni les mêmes problèmes à traiter.

¹ Les leviers dont il est question dans ce rapport peuvent être agronomiques, techniques, biologiques.

² L'IFT comptabilise le nombre de doses de référence utilisées par hectare au cours d'une campagne culturale. Pour un traitement phytosanitaire, l'IFT est le ratio entre la dose employée et la dose de référence du produit phytosanitaire utilisé pour une culture donnée(14). Il permet d'estimer le nombre de doses par hectare sur une campagne complète et se calcule de la manière suivante :

$$IFT = \frac{\text{dose de produit appliquée}}{\text{dose homologuée du produit}} * \frac{\text{surface traitée}}{\text{surface totale considérée}}$$

2.1.3. L'expérimentation en arboriculture

Dans le cas d'EcoPêche, on veut étudier l'effet de plusieurs pratiques combinées pour réduire les produits phytosanitaires. Il n'est donc pas ici possible de mettre en place un ensemble d'expérimentations suivant un plan factoriel classique. On s'intéresse aussi à la durabilité. Pour ce faire, il faut nécessairement adapter les pratiques au fil du temps.

L'arboriculture peut s'avérer être un sujet d'expérimentation complexe. D'une part, il n'est pas possible (comme cela pourrait être le cas pour des plantes annuelles) de réaliser les expérimentations dans un milieu où un ensemble de paramètres pourraient être contrôlés tels qu'une serre. D'autre part, l'arboriculture étant plus couteuse et complexe à gérer du fait de la pérennité des vergers et de l'espace nécessaire, il n'est pas toujours possible de réaliser des essais systèmes en respectant le concept d'une planification expérimentale (à moins de ne s'intéresser qu'à quelques facteurs). De plus, la rentabilité des vergers prenant parfois le pas sur l'expérimentation, les paramètres sont parfois modifiés d'une année sur l'autre ou en cours d'année. Il n'est d'ailleurs pas évident que les différents sites suivent les mêmes pratiques. Chaque site peut agir comme il le souhaite du moment que les règles de décisions mises en place dès le début sont respectées. Bien que les systèmes de cultures soient suivis dans le temps, certaines de leurs caractéristiques peuvent donc évoluer.

Les vergers, pour chaque site, sont soit côte à côte, soit plus éloignés selon les terrains disponibles lors de la mise en place des vergers. Ces expérimentations demandant de très grandes surfaces, il est possible que des hétérogénéités sur un même terrain influencent les résultats.

Il n'y a pas de répétitions de système sur chaque site.

Selon ce que l'on veut étudier, deux échelles peuvent être prises en compte : celle de l'arbre (avec des opérations spécifiques à chaque arbre, comme la taille) et celle du verger (avec des opérations plus globales comme la fertilisation et l'irrigation). Dans les cas de mesures sur les arbres, seul un échantillon d'arbres par système est mesuré et non l'ensemble des arbres du système.

2.2. Objectif du stage

Des synthèses annuelles par site ont déjà été effectuées. Cependant, l'exploitation des données du projet à des échelles multisites et pluriannuelles n'a encore jamais été effectuée. L'analyse de ce type d'expérimentation est donc difficile à aborder du fait des caractéristiques du projet, et on ne dispose pas de méthodes statistiques adaptées à des plantations expérimentales combinant des multiples facteurs, sans répétitions et avec des variables très hétérogènes.

L'objectif du stage est de proposer et mettre en œuvre une démarche statistique adaptée afin d'exploiter au mieux les données du projet EcoPêche. Il s'agira d'évaluer les différences de performances multicritères entre les systèmes de culture testés et d'identifier les principaux facteurs explicitant les résultats des systèmes. Déterminer les méthodes appropriées est capital pour pouvoir tirer des conclusions avec un maximum d'exactitude.

Les analyses ont pour but d'expliquer les performances des vergers de pêcheurs par la combinaison de leviers agronomiques. Il est également possible que ces combinaisons diffèrent selon les régions où sont implantés les vergers. La spécificité de ce projet est de tenir compte à la fois de l'aspect multisite et de l'aspect pluriannuel.

Concernant les performances que l'on cherche à expliquer, elles ne sont pas uniquement sur le plan agronomique, tels que les rendements obtenus, mais également techniques et économiques. On s'intéresse par exemple à l'aspect qualitatif des fruits, aux coûts de production et à la rentabilité des vergers. Les leviers peuvent être les différentes fréquences d'application de produits, l'irrigation ou la fertilisation apportée, l'utilisation de piégeage, etc. On portera un intérêt particulier à l'utilisation ou

non-utilisation d'une vingtaine de pratiques agricoles employées pour lutter contre les adventices¹, maladies ou ravageurs². L'IFT sera fréquemment employé car il est l'indicateur de suivi de l'utilisation des pesticides le plus synthétique et approprié.

3. Matériels et Méthodes

L'ensemble des analyses ont été effectuées sous le logiciel R (version 3.5.2)

3.1. Présentation des données

Cette partie du rapport aborde principalement les aspects de data-management qui ont été réalisés afin d'obtenir une base de données propre aux analyses que l'on souhaite effectuer. Les descriptions des données qui y sont faites ont surtout pour but d'en comprendre la structure.

3.1.1. Description des bases de données

Après avoir lu des documents relatifs au projet EcoPêche afin d'en comprendre le contexte, les bases de données relatives à ce projet m'ont été transmises.

Le fichier d'origine est un document au format Excel comportant 4 tables différentes dont on ne retiendra que deux (les deux autres étant des duplicatas de données présentées différemment ou des données qui ont sciemment été écartées des analyses car elles n'étaient pas jugées pertinentes).

La première table est divisée en 2 parties. La première partie est une description qualitative des systèmes³ (type de système, variété, porte-greffe, type de fruits, groupe de précocité, année de plantation, densité de plantation, forme fruitière). La seconde partie présente un ensemble d'indicateurs de performance ou de paramètres agronomiques, techniques ou économiques (rendement agricole, calibre des fruits, temps de travail, IFT, fertilisation, irrigation, efficience, coûts et rendement économique) en termes quantitatifs. Ces données sont présentées pour chaque année couvrant la période 2013-2018.

La seconde table contient la liste des leviers d'action par sites (de 2016 à 2017) ordonnée par type de bioagresseurs visés par ceux-ci. Après plusieurs discussions, une nouvelle table portant sur les leviers d'action agronomique m'a été transmise, car la précédente ne portait que sur deux des années de suivi et certains systèmes n'étaient pas décrits. La nouvelle base était plus complète et contenait les informations d'une année supplémentaire (soit des années 2016 à 2018). Le suivi des leviers n'ayant pas été effectué avant 2016, ces informations sont manquantes pour les premières années de suivi des vergers.

3.1.2. Description des données

La structure des données récoltées est importante à retenir pour comprendre les méthodes statistiques, les choix et les difficultés rencontrées.

¹ En agronomie, ce terme est aussi utilisé comme synonyme de « mauvaise herbe » (plante herbacée ou ligneuse indésirable à l'endroit où elle se trouve)(4).

² Organisme nuisible qui vit directement aux dépens de plantes ou de denrées en provoquant des dommages. Lorsque ceux-ci sont peu importants, on parle de déprédateur. Ce terme s'applique surtout aux animaux(4).

³ On fait ici référence aux différents systèmes expérimentaux du projet EcoPêche. On emploiera uniquement le terme « système » dans l'ensemble du rapport.

Chaque système pour une année donnée est décrit par une seule ligne. Les systèmes seront donc les individus statistiques de cette étude. Il n'y a pas de répétition d'un même système pour une année donnée.

Les données pour un système donné sont synthétisées sur l'ensemble du système pendant une année. Par exemple, le rendement commercialisable en frais (exprimé en tonne par hectare) représente un rendement spécifique (techniquement commercialisable) obtenu sur l'ensemble du verger correspondant à un système pour une année donnée, de même que l'IFT total représente la proportion de traitement répandu sur la totalité du système pour une année donnée.

Parmi l'ensemble des variables dont on dispose, certaines représentent les mêmes informations avec des niveaux de détail différents, et bon nombre de variables sont la somme d'autres. C'est le cas notamment avec les IFT. L'IFT total est la somme de plusieurs variables IFT parfois elles-mêmes regroupées selon le type d'IFT (vert ou chimique) et/ou sa finalité (herbicides, fongicide), ce qui fait un total de 14 variables pour décrire les IFT (Annexe 3).

La multiplicité de variables dont on dispose pour décrire des paramètres est déroutante car les différences tiennent du détail agronomique et il n'est pas facile de déterminer quelles sont les variables les plus appropriées à étudier.

3.2. Préparation des données

3.2.1. Importation et fusion des bases de données

Afin d'importer correctement les données, des fichiers ont été créés en abrégant les noms des variables afin que celles-ci soient plus aisément utilisables pour les codes des analyses à venir. Une fois les fichiers importés dans leur totalité, une suppression des variables ne comportant aucune information a été effectuée dans la table des indicateurs de performances. La fusion des fichiers a été effectuée en se basant sur les sites, les types de système et l'année de suivi (un changement temporaire d'intitulé des systèmes a été nécessaire pour la bonne fusion des informations).

3.2.2. Choix des sites et de la période

Bien qu'il n'ait été question que de 7 sites jusque-là, la base en contenait initialement 10. Trois furent écartés d'emblée des analyses. Il a été déterminé auparavant que les sites « SERFEL 3 », « Bordeaux », et « Avignon 2 » ne correspondaient pas en terme d'attente et d'expérimentation, et que leur prise en compte risquerait de biaiser les analyses.

En raison des données manquantes lors de la phase de préparation des vergers (« jeunes vergers ») et des indicateurs ne pouvant pas être relevés à ce stade, on ne considèrera que les données portant sur la phase de production des systèmes.

3.2.3. Uniformisation des types de système

Selon les sites, les systèmes ne sont pas clairement classés selon le type de système (REF, ECO et BIO). Il était donc nécessaire de re-catégoriser les différents systèmes. Si cela était évident pour la plupart des systèmes, quelques-uns n'étaient pas clairement nommés et changeaient parfois d'appellation d'une année sur l'autre en raison de modifications expérimentales, ce qui a été assez long à gérer.

3.2.4. Ajout d'identifiants et variable temporelle

Avant cela, il est nécessaire pour pouvoir distinguer les systèmes lorsque plusieurs systèmes d'un même type (REF, ECO ou BIO) sont situés sur le même site (la combinaison des variables site

et type de système n'est pas suffisante pour identifier chaque système). Un identifiant des systèmes est alors mis en place (

Annexe 1). Cet identifiant sera plus simple à utiliser pour déterminer l'évolution des indicateurs de performance dans le temps. Une fois l'identifiant des systèmes mis en place, on peut renommer proprement les types de système en REF, ECO ou BIO.

Un identifiant pour chaque combinaison de système-année a également été ajouté afin de pouvoir, au besoin, retrouver ou rectifier des données précises.

En prévision des analyses utilisant les modèles mixtes, il est nécessaire de disposer d'une variable représentant le temps depuis le début du suivi des systèmes. Il a été choisi de ne s'intéresser au verger de pêcheurs uniquement lorsque ceux-ci sont en phase de production (ce qui correspond à un verger planté depuis 3 ou 4 ans). Si l'on dispose d'informations sur les systèmes lors de leur phase de formation (jeune verger), la majorité des indicateurs sont manquants, car d'une part, certains ne peuvent être relevés, ou d'autre part, les vergers n'étaient pas encore suivis lors de cette période. Pour la suite des analyses, on peut considérer que les vergers sont étudiés à partir de leur première année de production (considérée comme l'instant t_0).

Les données sont écartées dès qu'un événement climatique important survient. En effet, celui-ci perturbant la production et les indicateurs qui y sont liés, il n'a pas été jugé convenable de prendre en compte les systèmes concernés et n'assurant pas le potentiel de production attendu lors de ces années (celles-ci diffèrent d'un site à l'autre). Les incidents climatiques sont une des principales causes de la baisse de production de pêches au cours des dernières années(15).

3.2.5. Recodage des leviers qualitatifs

Les leviers issus de la seconde table sont des variables qualitatives décrites initialement selon trois modalités selon l'efficacité attendue de ces leviers sur les différents systèmes. La notion d'efficacité a été évaluée par un groupe d'experts. Néanmoins, pour éviter d'avoir un jugement sur l'efficacité des leviers, il a été choisi de regrouper les catégories afin de tenir compte uniquement de l'utilisation sans a priori sur leur efficacité.

3.3. Méthodes d'analyses utilisées

3.3.1. Analyse descriptive

Les analyses descriptives seront stratifiées afin d'obtenir une première approche des données sans être perturbé par les corrélations liées aux systèmes, sites ou années. Ces aspects ne peuvent cependant pas être tous représentés en même temps.

En regroupant les informations par site ou année, quelques tests statistiques non-paramétriques sont effectués (Kruskal-Wallis, Wilcoxon). Face aux faibles effectifs, les tests adaptés pour deux variables qualitatives ne seront pas présentés ou effectués, leur application nécessitant la réalisation de nombreux tests du χ^2 de Yates ou exact de Fisher sur des tableaux de contingence 2x2.

Tel qu'il était prévu initialement, nous nous sommes concentrés surtout sur des méthodes d'analyses factorielles ou longitudinales. Néanmoins, les données dont nous disposons sont problématiques pour les analyses que nous souhaitons réaliser.

3.3.2. Analyses factorielles

S'intéressant à des performances multicritères et disposant d'un vaste ensemble de paramètres agronomiques, visualiser en même temps un ensemble de variables à expliquer et explicatives est ici pertinent. Les méthodes d'analyses factorielles offrent cette possibilité en permettant de résumer et

de visualiser l'information contenue dans des données multivariées. Bien que ces méthodes n'apportent pas de réponse statistique concrète et laissent une interprétation assez subjective, elles sont couramment employées, notamment dans les analyses d'enquête(16,17).

Mais les analyses factorielles ne sont pas adaptées aux données groupées ou répétées. En effet, ces méthodes ne prennent pas en compte les corrélations intra-sujets.

Dans un premier temps, ces méthodes ont donc été mises en place pour chacune des années de suivi (une analyse par année¹). Mais le nombre de cas dont nous disposons chaque année est variable et parfois très faible (par exemple, seuls 4 systèmes en 2013).

Lorsque l'on réalise des analyses factorielles et que l'on a très peu d'individus, l'analyse peut être biaisée car la contribution de chaque individu sera trop importante (surreprésentation). Par ailleurs, l'interprétation pourrait être plus complexe car on ne pourra pas forcément distinguer des individus ayant des caractéristiques similaires. Il est conseillé de disposer d'un nombre élevé d'individus n au moins trois fois plus élevé que le nombre de variables p utilisées pour ces analyses, faute de quoi les résultats seront peu stables. Ce qui n'est évidemment pas le cas ici si on réalise des analyses année par année.

Afin d'obtenir des analyses factorielles dont les résultats seront plus robustes, il a volontairement été choisi de ne pas tenir compte de l'aspect pluriannuel des données et de prendre en compte l'ensemble des systèmes-années dans les analyses. De plus, les paramètres pouvant changer d'une année à l'autre, on peut considérer chaque système-année comme étant un individu propre. Les analyses factorielles ainsi construites seront plus adéquates car l'on disposera de plus d'individus et l'on pourra mieux visualiser des groupes de système ayant les mêmes caractéristiques. Il est vrai que de cette manière, on néglige l'aspect temporel, mais un effet site pouvant également rentrer en compte, il sera de toute manière obligatoire de faire des concessions lors de l'analyse.

3.3.2.1. *Analyse en composantes principales (ACP)*

L'ACP est une méthode statistique servant à explorer des données multivariées lorsque celles-ci sont quantitatives. Elle permet l'étude des ressemblances et des différences entre individus du point de vue de l'ensemble des variables, ainsi que des liaisons entre les variables. Les données vont être représentées sur des axes (ou composantes) qui sont des variables synthétiques créées à partir d'un ensemble de variables corrélées. Par défaut, les données sont centrées et réduites. Il est également possible de représenter la contribution des variables à la formation de l'axe, ainsi que la qualité de représentation des variables(16,17).

Le résultat d'une ACP se visualise par le graphique des variables et celui des individus mais ces deux graphiques peuvent être regroupés en un seul.

3.3.2.2. *Analyse des correspondances multiples (ACM)*

Les données multivariées qualitatives sont analysées via des ACM dont le principe est le même que pour une ACP (16,17). Le résultat d'une ACM se présente en trois parties.

Le graphique des individus rend possible l'étude des ressemblances entre individus du point de vue de l'ensemble des variables qualitatives. Le graphique des variables dresse le bilan des liaisons entre variables et résume l'information contenue dans les composantes retenues. Le graphique des

¹ Initialement, il a été choisi de stratifier par années afin de garder un minimum de contrôle sur les facteurs de confusion climatiques. Il aurait été intéressant de stratifier selon l'année de récolte (temps t employé dans les analyses longitudinales).

modalités permet de comprendre les proximités et les éloignements entre les modalités des différentes variables étudiées. De ce fait, on pourra distinguer si les modalités de différentes variables sont communes ou non à un groupe d'individus. Il faut cependant garder à l'esprit que les modalités trop rares auront une importance trop forte (inertie élevée), et risquent de contribuer exagérément à la formation des axes. Il faut donc les gérer, soit par regroupement de modalités ou par la répartition d'individus dans d'autres modalités (ventilation). Dans notre cas, soit la première option fut choisie, soit, ayant un grand nombre de variables qualitatives binaires, la variable fut écartée des analyses.

La recherche du critère de Kaiser ou celui de Cattell pour déterminer le nombre d'axes à prendre en compte est effectuée pour chaque ACP et ACM. Mais il a parfois été choisi sciemment un nombre d'axes selon le pourcentage d'inertie¹ que l'on souhaite retenir ou l'interprétation que l'on peut en faire sans retenir ces critères précédemment cités si ceux-ci indiquait un nombre d'axes trop important. Concernant le pourcentage d'inertie, on cherchera idéalement à conserver près de 75% de l'inertie totale.

3.3.2.3. *Classification ascendante hiérarchique (CAH)*

Cette méthode a pour objectif de répartir ces individus dans un certain nombre de classes selon la méthode ou le choix effectué. En se basant sur une matrice des distances représentant les dissimilarités entre chaque individu pris deux à deux, la CAH produit un arbre de classification (ou dendrogramme). Comme le font les ACP ou ACM, ce dernier permet de visualiser les proximités entre individus. En découpant cet arbre à une certaine hauteur choisie, on produira une partition divisant l'ensemble des individus en classe (16,17).

Dans les cas présentés ici, la distance considérée pour les dissimilarités est la distance euclidienne (qui est souvent la distance employée par défaut). La méthode de Ward est également appliquée afin de regrouper les éléments de manière à optimiser la qualité de la partition obtenue. Pour rappel, la méthode de Ward regroupe des individus ou ensemble d'individus en minimisant l'inertie intra-classe à chaque agrégation, bien que celle-ci augmente la perte d'inertie interclasse.

Dans les premiers temps, les partitions utilisées sont celles proposées par défaut sous R dont le critère de sélection est le maximum de la perte d'inertie-intra relative en augmentant d'une valeur le nombre de classe. Mais, afin d'obtenir des partitions optimales, le critère de Ward fut retenu. Dans un souci d'interprétation des clusters, le critère pour la réalisation des partitions reste à définir selon les cas. Parfois, nous avons choisi le nombre de clusters sans tenir compte de ces critères si les groupes étaient visuellement détectables et plus facilement interprétables, tandis que le critère de Ward a été parfois le critère de partitionnement retenu. Cette méthode a finalement été peu utilisée car vite supplantée par la classification hiérarchique sur composantes principales, mais il reste important d'en comprendre le concept pour utiliser la méthode suivante.

3.3.2.4. *Classification Hiérarchique sur Composantes Principales (HCPC)*

Si l'analyse factorielle et la classification ascendante hiérarchique partagent le même objectif, celles-ci diffèrent dans les méthodes et représentations obtenues. On peut combiner ces deux méthodes afin d'enrichir et de renforcer l'analyse exploratoire (16,18).

La classification hiérarchique sur composantes principales (communément abrégé HCPC d'après son acronyme anglais) permet de construire un arbre de classification. Lorsque cet arbre est construit, il est proposé de couper cet arbre afin de construire des classes homogènes suivant le même principe

¹ L'inertie représente l'information du tableau de données contenue dans les différents axes.

que la classification hiérarchique classique, à la différence que ces classes sont ensuite consolidées par le partitionnement en k-means. En se basant sur une partition initiale, la méthode des k-means réalise une succession de calculs des centres de gravité des différentes classes et une réaffectation des individus à la classe la plus proche jusqu'à ce que celle-ci ne soit plus modifiée (convergence). Bien que l'aspect hiérarchique soit alors perdu (il n'y a parfois plus de correspondance entre les classes et l'arbre), les classes constituées sont plus homogènes. Cette dernière option est utilisée dans les cas présentés dans ce rapport, mais elle peut également être désactivée.

Sachant que la CAH nécessite des données quantitatives, cette approche est d'autant plus utile lors de la prise en compte de données qualitatives car la classification pourra être effectuée sur les coordonnées des individus fournis par l'ACP ou l'ACM.

Une fois que les classes d'individus sont formées, il est important de savoir comment les caractériser. Les classes peuvent être décrites par leur parangon (l'individu le plus proche du centre de gravité d'une classe) qui est l'individu le plus représentatif du comportement moyen d'une classe. Il est également possible de déterminer les variables qui caractérisent le mieux les classes.

3.3.2.5. *Autres méthodes*

Dans le cadre de l'analyse des données du projet EcoPêche, d'autres méthodes ont été abordées. Celles-ci ont été utilisées dans une moindre mesure par rapport aux précédentes méthodes citées. Même si ces méthodes ont été peu utilisées durant ce stage, il reste cependant important de les citer car elles pourront éventuellement être réutilisées pour analyser des données similaires.

4. *Analyse Factorielle des Données Mixtes (AFDM)*

L'analyse factorielle des données mixtes (AFDM) est une méthode d'analyse factorielle permettant d'inclure à la fois des variables quantitatives et qualitatives d'une même analyse (19). L'ensemble des variables est normé et centré afin de prendre en compte l'ensemble des variables pour le calcul des distances entre individus, variables quantitatives et qualitatives ainsi que les modalités de celles-ci. Les résultats se présentent comme ceux qu'une ACP et d'une ACM (les variables quantitatives sont représentées dans un cercle de corrélation, tandis que les variables qualitatives et leurs modalités font l'objet de deux autres graphiques).

5. *Analyse en Composantes Principales de Variables Instrumentales (ACPVI)*

Le concept des ACPVI, aussi appelés analyse des redondances (20), consiste à penser que chaque variable à expliquer Y est décrite par un modèle basé sur les variables explicatives X. On cherche alors à expliquer une ACP des variables Y par une ACP des variables X. L'ACPVI peut être vue comme une ACP des prédictions du tableau Y obtenues par régressions multiples sur les variables de X. Cette méthode est assez intéressante pour comparer et expliquer un ensemble de variables Y par X mais sa réalisation et son interprétation restent complexes, raison pour laquelle cette méthode ne se sera que peu abordée dans ce rapport.

5.1.1. *Analyse de données longitudinales*

Afin de prendre en compte l'aspect longitudinal des données, la réalisation de modèles mixtes est une méthode adéquate. L'ensemble des performances nous intéressant étant des variables quantitatives, il s'agit précisément de réaliser des modèles linéaires mixtes. Pour se faire, on considèrera que nos données suivent une loi gaussienne¹.

¹ Ces hypothèses seront vérifiées par l'étude des résidus issus des modèles réalisés.

Par rapport à des modèles linéaires classiques, il est possible d'ajouter dans le modèle des effets aléatoires. L'intercept aléatoire est un intercept spécifique à chaque sujet (représentant l'effet sujet) introduit dans le modèle afin de tenir compte de la corrélation intra-sujet tout en restant dans le cadre des modèles de régression ordinaire. Cet intercept n'a pas de valeur fixe, il est aléatoire, supposé gaussien, d'espérance nulle et de variance σ_b^2 . L'intercept aléatoire permet d'inclure un effet propre à chaque groupe ou individu. D'autres effets aléatoires peuvent être introduits dans les modèles, notamment des pentes aléatoires pour les modèles sur données longitudinales (les pentes impliqueront une évolution dans le temps propre à chaque individu).

Rappel de la formulation d'un modèle mixte standard :

$$Y_{ij} = X_{ij}^T \beta + Z_{ij}^T b_i + \epsilon_{ij} \text{ Avec } \epsilon_{ij} \sim N(0, \sigma^2) \text{ et } b_i \sim N(0, B)$$

Concernant l'application de modèles mixtes sur les données du projet EcoPêche, nous faisons face à des problèmes de convergence. Suite à cela, certains modèles qu'il serait intéressant d'étudier n'arriveront donc pas à être estimés. Il est donc nécessaire de faire des choix dans la structure des modèles.

Considérant que nous avons un effectif d'étude assez petit (19 systèmes), nous pourrions choisir d'employer le maximum de la vraisemblance restreinte (REML) plutôt que le maximum de vraisemblance (ML). Néanmoins, il n'est pas possible de comparer des modèles basés sur le maximum de la vraisemblance restreinte s'ils ne contiennent pas les mêmes effets fixes. La majorité des modèles seront donc réalisés avec le maximum de vraisemblance afin de pouvoir comparer les modèles et chercher les plus adaptés à notre cas. Par ailleurs, dans de nombreux modèles testés par la suite, les problèmes de convergence rencontrés surviennent quelque soit la méthode choisie (ML que REML), et ce, généralement lors de l'ajout d'une pente aléatoire. En raison de ces problèmes de convergence, il sera parfois nécessaire de se contenter d'un intercept aléatoire, sans pouvoir tester l'apport de pentes aléatoires. Nous n'utiliserons également peu de variables explicatives par modèle.

5.2. Gestion des données manquantes

On supposera que les données manquantes sont MCAR¹. Dans le cas où les variables comportaient des données manquantes, celles-ci ont été gérées par imputation multiple.

Dans un maximum de cas, on limitera l'usage d'imputation des données manquantes. Seules les variables ayant un intérêt particulier et une faible proportion de données manquantes seront imputées. Face aux variables incomplètes apportant des informations détaillées, on leur préférera des variables apportant la même information avec moins de détail si celles-ci sont complètes (par exemple, on préférera la variable *ift_vert_total* aux variables incomplètes *ift_vert_insect* ou *ift_vert_fongi*), ou bien elles seront écartées des analyses si elles ne sont pas primordiales. Concernant les IFT, on considèrera que les données manquantes correspondent à des cas où l'IFT n'était pas utilisé donc peuvent a priori être remplacés par 0.

¹ Une donnée est MCAR (missing completely at random) est manquante de façon complètement aléatoire si la probabilité d'absence est la même pour toutes les observations.

5.3. Aspect programmation

5.3.1. Package et fonction R utilisés

Ces analyses factorielles ont été réalisées avec les fonctions *PCA*, *MCA*, *FAMD*, *HCPC* du package FactoMineR et la fonction *PCAIV* du package ade4(21), tandis que les modèles pour données longitudinales ou groupées ont été effectués avec la fonction *lme* du package nlme.

L'imputation multiple a été effectuée à l'aide du package missMDA pour les analyses factorielles et du package mice pour les analyses longitudinales ou groupées.

5.3.2. Optimisation des graphiques sous R

Afin d'optimiser l'aspect visuel sur un même graphique concernant les analyses factorielles et cela dans de nombreux cas :

- Les systèmes sont représentés sous une appellation correspondant à l'identifiant du système et l'année correspondante. Par exemple, la valeur du système dont l'identifiant est « 1 » (système de référence de l'INRA d'Avignon) pour l'année 2018 est nommé « 1-18 ».
- Les variables sont colorées selon leur contribution à la formation des axes (pour les ACP uniquement).
- Les individus sont colorés selon le type de système (REF, ECO ou BIO) auquel ils appartiennent.
- Des ellipses (avec un intervalle de confiance à 95%) des individus correspondants sont ajoutées pour chacun des systèmes ou clusters.

La majorité des graphiques ont été effectués avec la fonction *ggplot* après avoir récupéré les coordonnées, et autres valeurs importantes à représenter graphiquement. Cette démarche a permis d'obtenir des graphiques plus facilement manipulables et plus ergonomiques que les graphiques proposés par défaut.

6. Résultats

Cette partie du rapport est divisée en deux parties. La première partie sera descriptive et permettra d'avoir un aperçu global des données. Dans une seconde partie, on présentera les résultats des analyses factorielles, longitudinales ou répétées.

6.1. Analyses descriptives

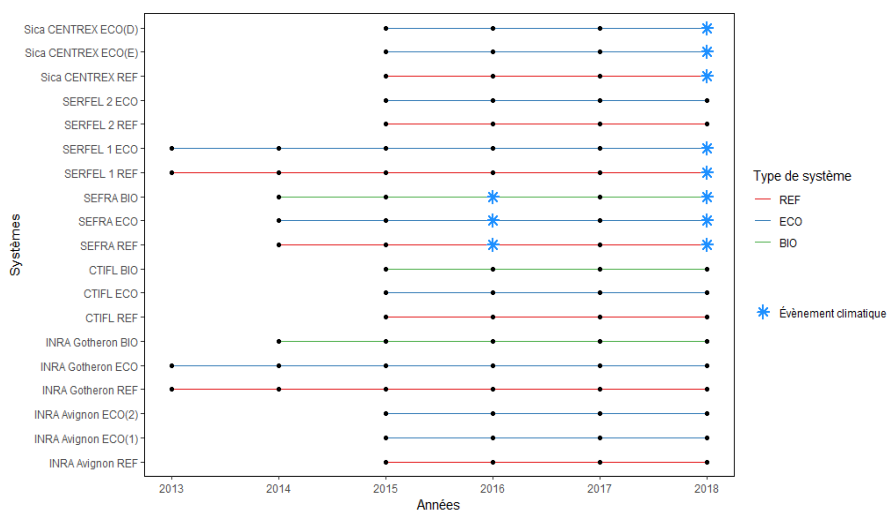
Il est important de disposer d'une vision globale des données. Ainsi, on pourra distinguer des relations ou effets de confusion qui pourraient influencer les résultats à venir.

6.1.1. Caractéristiques générales des systèmes

Si on ne tient compte que des caractéristiques générales (structure des vergers), les caractéristiques des systèmes sont très similaires pour un même site, mais sont très différentes d'un site à l'autre (Annexe 4). Par exemple, les variétés diffèrent pour chaque site. De plus, les variétés entre les systèmes de INRA Gotheron ne sont pas les mêmes. Chaque site produit un seul type de fruit (nectarine, pêche), la densité de plantation et la forme fruitière diffèrent souvent d'un site à l'autre mais peuvent parfois être les mêmes.

6.1.2. Suivi des systèmes

Graphique 1 : Suivi des systèmes dans le temps (années)



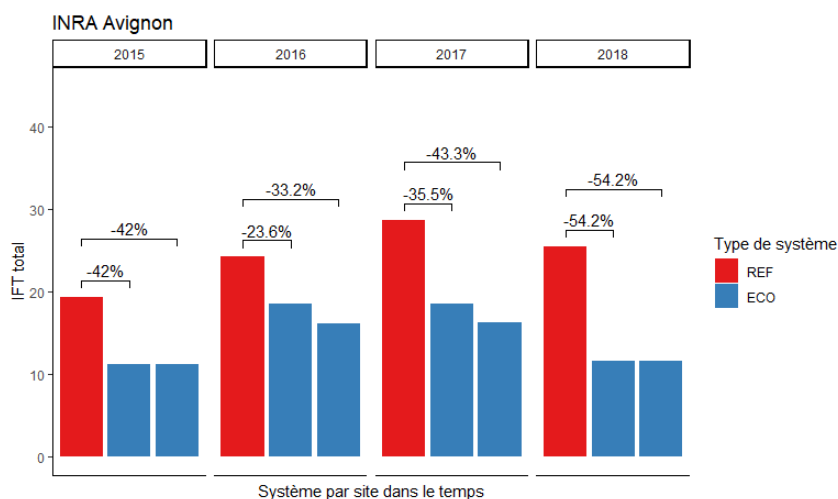
Durant le suivi, 3 sites ont subi des événements climatiques ayant entraîné la perte de la récolte et donc d'informations pour nos analyses. Les sites touchés étaient Sica Centrex, Serfel 1 en 2018, ainsi que Sefra en 2016 et 2018. La durée de suivi est assez courte. En tenant compte des événements climatiques, les systèmes ont un suivi moyen de 4,05 ans (6 sont suivis pendant 3 ans, 8 pendant 4 ans, 3 pendant 5 ans et 2 pendant 6 ans). Pour rappel, on ne considère que la phase de production des vergers.

6.1.3. Réduction des intrants

Si la réduction des produits phytosanitaires est l'un des enjeux de ce projet, la réduction de la dose d'IFT visé n'est pas la même d'un site à l'autre. Pour exemple, les sites INRA Avignon, INRA Gotheron, CTIFL, et SERFEL 2 visaient une réduction de 50% des intrants, tandis que SERFEL 1 et SEFRA prévoyaient une réduction de 30%.

Même en tenant compte de cela, peu de systèmes arrivent à atteindre les objectifs qu'ils s'étaient fixés, ou du moins, cet objectif n'est pas atteint sur l'ensemble des années (Annexe 5). Lors des premières années, la réduction est beaucoup moins effective que prévue, mais a tendance à être plus proche des objectifs visés après quelques années.

Graphique 2 : Différence annuelle d'IFT total entre les systèmes par rapport au système de référence (site d'Avignon)



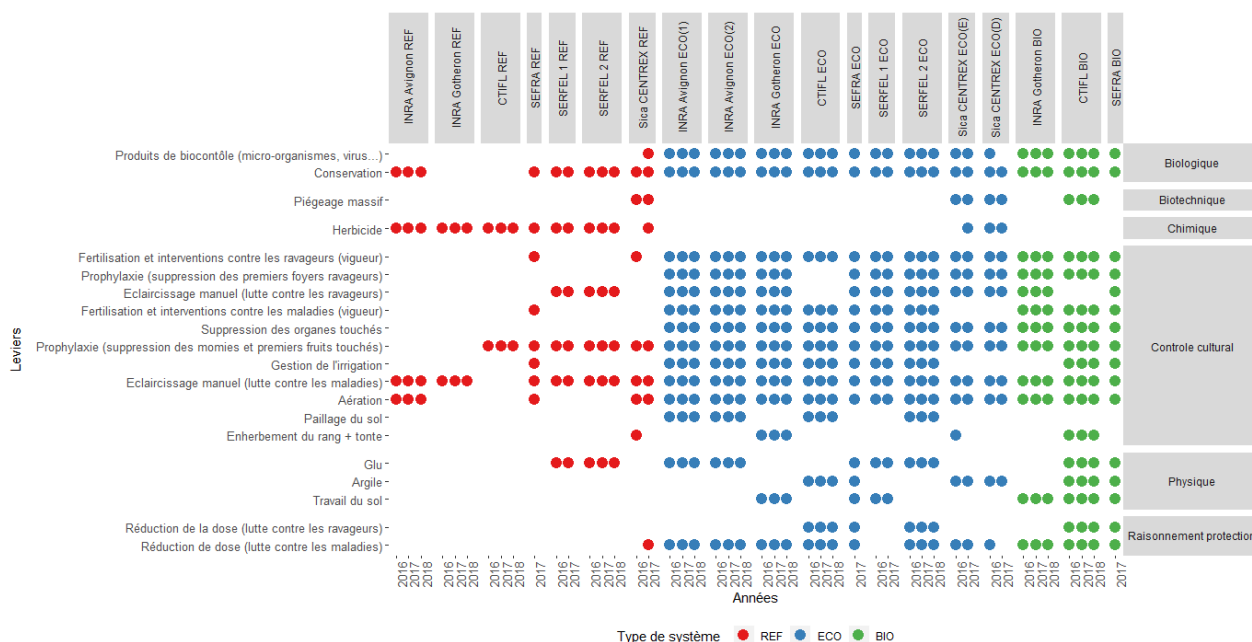
6.1.4. Utilisation des leviers

Pour information, des leviers ont le même intitulé mais n'ont pas la même destination. Certains luttent contre les maladies, tandis que d'autres luttent contre les ravageurs, ou encore contre les adventices.

Après le recodage des leviers en variables binaires (utilisé ou non), plusieurs leviers n'apportent plus de variabilité : ils sont soit toujours, soit jamais utilisés. Neuf leviers¹ sont ainsi écartés. Deux autres leviers² ont finalement été écartés car d'une part, l'utilisation de ces leviers était discutable (méthodes peu employées et de ce fait, on ne leur accorde pas trop d'intérêt) et d'autre part, leur utilisation étant rare dans notre étude, ces leviers risquaient de prendre trop d'importance lors des analyses multidimensionnelles.

Après la création d'un graphique pour synthétiser l'utilisation des différents leviers retenus (Graphique 3), on cherche à distinguer les leviers qui caractérisent les systèmes ou type de systèmes. Il apparaît que ces leviers sont souvent réutilisés d'une année sur l'autre par les différents systèmes.

Graphique 3 : Utilisation des leviers par système et années, groupés selon le type de levier



En comptant le nombre de leviers utilisés selon leur type³ (biologique, biotechnique, chimique, etc.), il est possible de distinguer statistiquement les types de leviers qui caractérisent les types de systèmes (Kruskal-Wallis ou Wilcoxon). Nous préférons ici cette approche plus synthétique à l'étude des leviers un par un. Quelle que soit l'année (2016 à 2018), le nombre de leviers chimiques utilisés est significativement plus élevé chez les systèmes de référence que par les deux autres types de systèmes. D'autre part, les systèmes ECO et BIO utilisent significativement plus de leviers biologique, physique,

¹ L'atténuation - conduite des arbres (Taille en vert / ébourrage), l'OAD + raisonnement (lutte contre les maladies), l'OAD + raisonnement (lutte contre les ravageurs), et la confusion sexuelle sont toujours utilisés, tandis que le lâcher inoculatif, le lâcher inondatif, les produits de biocontrôle alternatifs, le contrôle génétique (lutte contre les ravageurs), et les filets anti-insectes ne sont jamais utilisés.

² Produits de biocontrôle (lutte contre les maladies) et contrôle génétique (lutte contre les maladies)

³ Le regroupement des leviers selon leur type a été mis en place assez tardivement au cours du stage.

de contrôle cultural, et de raisonnement. Les systèmes ECO et BIO ne présentent pas de différences significatives selon les types de leviers utilisés. Le nombre de leviers biotechniques ne diffèrent pas selon le type de système (Annexe 6).

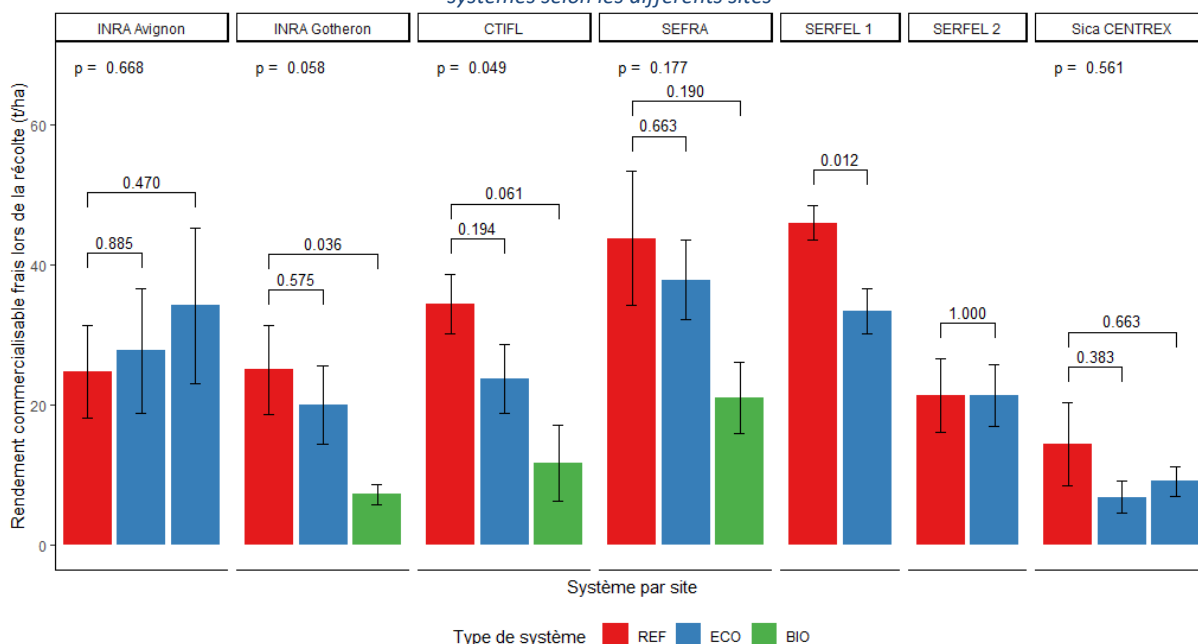
Les systèmes de référence emploient significativement moins de leviers que les systèmes ECO ou BIO ($p < 0.001$ dans les deux cas), avec une moyenne de 4.5 leviers (contre 13.7 pour les systèmes ECO et 14.7 pour les systèmes BIO) sur les 3 années étudiées. Le nombre de leviers ne diffère pas significativement entre les systèmes ECO et BIO ($p = 0.480$).

6.1.5. Comparaison des performances selon les systèmes

Dans l'optique de distinguer d'éventuelles différences de performances entre les systèmes et n'ayant qu'une valeur par système et par année, il a été choisi de considérer les années comme de simples répétitions des systèmes, ce qui permet de comparer statistiquement les systèmes entre eux pour chaque site.

Globalement, les rendements frais commercialisables (t/ha) des systèmes ECO sont plus faibles que leur référence, à part ceux de l'INRA d'Avignon, mais il n'y a pas de différences significatives entre les systèmes ECO et les systèmes REF leur correspondant, à l'exception du système ECO de SERFEL 1 qui est significativement plus faible que sa référence ($p = 0.012$). Les systèmes BIO ont des performances bien moins importantes (l'un d'eux est significativement plus faible) (Graphique 4).

Graphique 4 : Comparaisons des rendements frais commercialisables (t/ha) entre les systèmes de références et les autres systèmes selon les différents sites



Des tests de Kruskal-Wallis sont effectués pour comparer l'ensemble des systèmes d'un même site, et des tests de rangs de Wilcoxon sont employés pour comparer chaque système au système de référence correspondant.

Concernant les coûts de production (par kg/fruit en euros), la majorité des systèmes ECO et BIO sont plus coûteux. Néanmoins, le système BIO de la CTIFL est le seul système qui soit significativement différent (ici, plus élevé) que le système de référence ($p = 0.030$) (Annexe 7). En termes de marge partielle¹, il n'y a aucune différence significative entre les types de systèmes selon leur site respectif

¹ Marge partielle = Chiffre d'affaire - Coûts de production

(Annexe 8). Il est cependant à noter que sept des systèmes étudiés ont eu une première ou une deuxième année de récolte économiquement non rentable (production à perte).

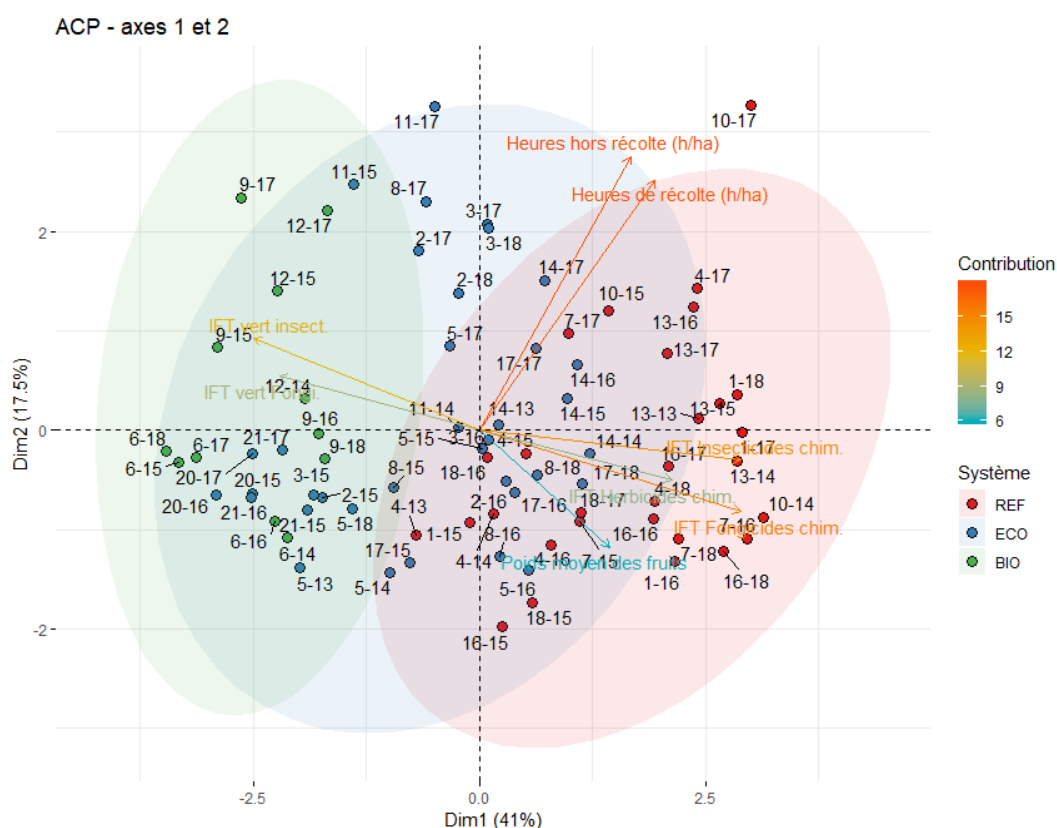
6.2. Analyses multidimensionnelles

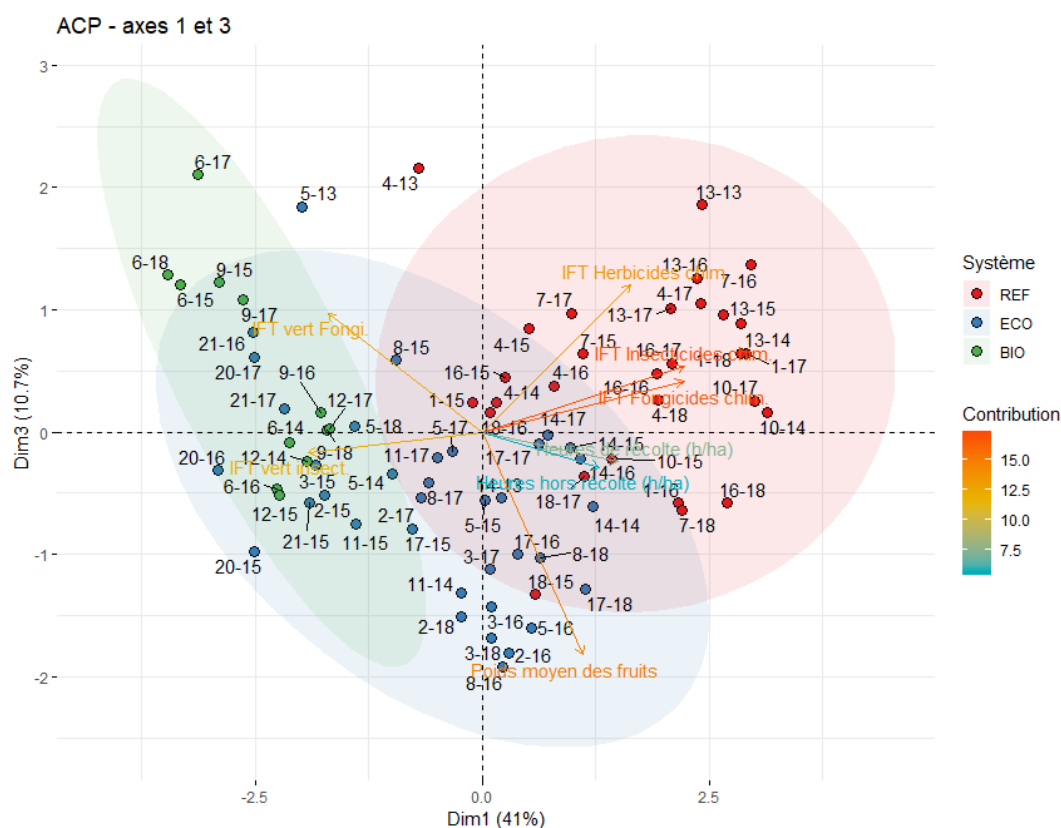
Dans l'optique de distinguer des corrélations et des groupes d'individus semblables, nous avons choisi de ne prendre que quelques variables en compte à la fois. Après une sélection de variables sur les conseils de mes encadrants, sur l'importance apparente de certaines variables synthétiques et le rejet de variables avec une forte proportion de données manquantes, des analyses factorielles ont été effectuées.

6.2.1. Quelques indicateurs de performances et variables agronomiques

Dans un premier temps, nous nous sommes intéressés à différents IFT ainsi qu'à quelques performances. Dans ce cas précis, il s'agit de deux variables sur le temps de travail effectué sur les systèmes et du poids moyen des fruits récoltés. Ces performances peuvent également être considérées comme des variables explicatives car d'autres variables plus synthétiques (agronomiques et économiques) vont dépendre de celles-ci.

Graphique 5 : ACP des IFT, heures de travail et poids moyen des fruits





La prise en compte de ces 3 axes permet d'obtenir près de 70% de l'inertie totale (contre près de 60% avec les 2 premiers axes préconisés par les critères de Kaiser et de Cattell)(Annexe 9).

Les IFT sont fortement corrélés à l'axe 1. On peut distinguer les IFT chimiques (à droite) des IFT verts (à gauche). Les heures de récolte et hors-récolte sont assez corrélées avec l'axe 2. Le poids moyen des fruits est corrélé avec l'axe 3. Les 2 premières dimensions permettent de distinguer clairement les systèmes BIO et REF. Les systèmes BIO emploient de plus grandes quantités d'IFT vert (fongicide ou insecticide) que les systèmes REF, et inversement en ce qui concerne les IFT chimiques (insecticides, fongicide ou herbicide). Les heures de récolte et hors-récolte ne semblent pas particulièrement associées à un type de système, les individus étant en haut du graphique (coordonnées élevées sur l'axe 2) n'appartenant pas à un seul type de système. On note par exemple que les systèmes du site SEFRA (n°10 et 11) ont nécessité plus d'heures de travail que d'autres systèmes en 2017. Si cette analyse permet de distinguer clairement les systèmes REF et les systèmes BIO, les caractéristiques des systèmes ECO paraissent plus hétérogènes. Ils ne se regroupent pas de façon aussi concise et sont parfois semblables à des systèmes BIO ou à des REF. En étudiant l'axe 3, il est cependant possible de noter que les systèmes ECO ont des poids moyens de fruits plus élevés que les 2 autres types de système.

En réalisant une HCPC sur cette ACP, on distingue 3 classes d'individus (Annexe 10, Annexe 11 et Annexe 12). Le premier, caractérisé par une proportion significativement plus élevée de système BIO et une absence de système REF, se distingue par des IFT verts élevés ainsi que des heures de récolte et des poids de fruits moins élevés. Le second, majoritairement composé de système ECO, se caractérise par l'utilisation d'IFT vert mais également par les IFT herbicides chimiques. Le troisième cluster, dont la proportion de système REF est significativement plus importante, est décrit par des IFT chimiques, des heures de récolte et des poids de fruits élevés. On retrouve ici les principales déductions que l'on obtenait avec l'ACP. Chaque cluster étant plus ou moins associé à un type de

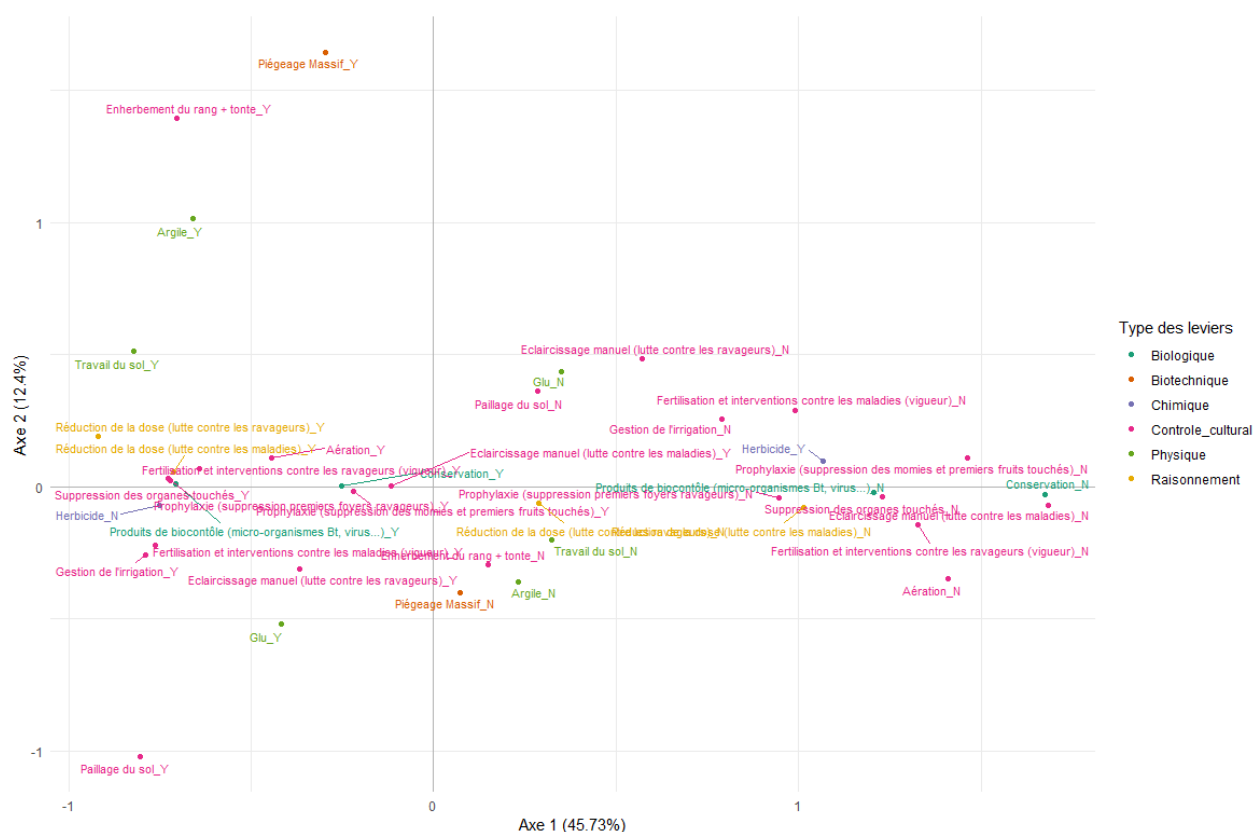
système, il est plus facile de décrire ceux-ci. Il est possible de noter, en se basant sur les parangons, que les systèmes ECO de l'INRA d'Avignon semblent proches des systèmes BIO. L'usage d'herbicide chimique caractérisant le deuxième cluster paraît être dû aux systèmes ECO de Sica Centrex qui en utilisent contrairement aux autres systèmes ECO.

Par la suite, une ACP portant sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique¹, le pourcentage de perte et le poids moyen des fruits est effectuée (Annexe 13). Il est possible de déduire de celle-ci ainsi que de l'analyse des clusters obtenus après une HCPC (Annexe 14, Annexe 15 et Annexe 16) que les systèmes BIO ont un pourcentage de perte plus élevé que les deux autres systèmes, et qu'entre les systèmes ECO et REF, deux groupes se distinguent. Certains ont des indices réfractométriques plus importants et des rendements moins élevés, et inversement.

6.2.2. Leviers qualitatifs

Afin d'avoir une vision plus détaillée des leviers et de distinguer ceux qui caractérisent les types de systèmes, une ACM est réalisée sur l'ensemble des leviers qualitatifs dont nous disposons. Ceux-ci sont colorés selon leur type. Il faut garder à l'esprit que les analyses impliquant les leviers ne concernent que les années 2016 à 2018.

Graphique 6 : ACM des leviers - graphiques des modalités

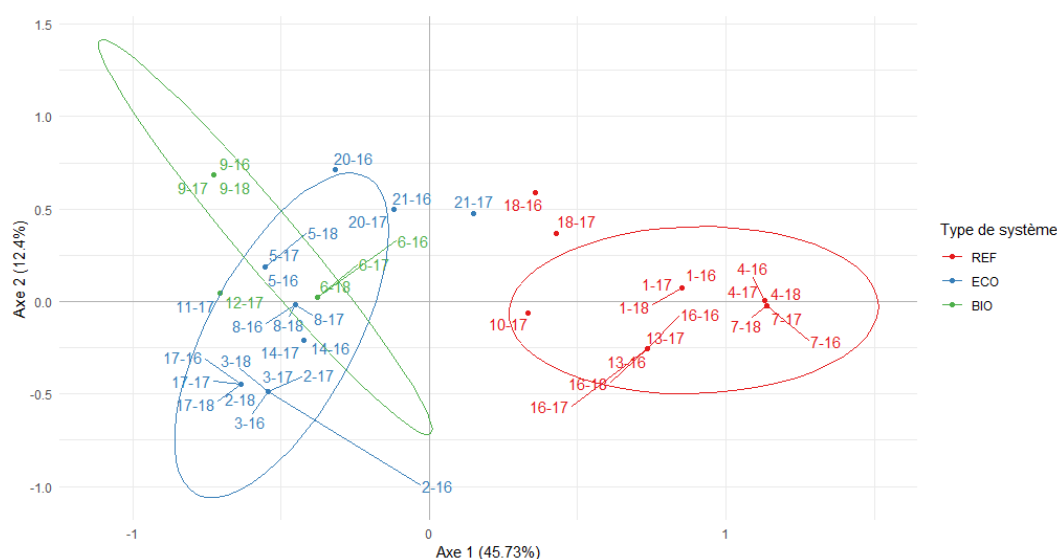


¹ L'indice réfractométrique représente la fraction de saccharose des fruits mesurée en degrés Brix (°B ou °Bx). Cette mesure permet ici d'estimer une valeur gustative des fruits. Mais celle-ci est très souvent négligée par les normes de commercialisation actuelles, celles-ci étant principalement axées sur les calibres des fruits.

L'inertie totale expliquée est de près de 60% avec les 2 premiers axes. Dans un souci d'interprétation des sorties, seuls ces 2 axes seront présentés. Sur le graphique des variables (Annexe 17), on distingue que le piégeage massif est fortement corrélé à l'axe 2, l'enherbement, le paillage au sol et l'argile sont également corrélés sur cet axe. L'axe 1 est fortement corrélé avec la vigueur, l'herbicide, et la suppression des organes touchés.

Le graphique des modalités permet de distinguer deux types de profils notamment via la première composante. A droite du premier axe, le profil est décrit par l'usage des herbicides et le non-usage de l'ensemble des autres leviers. A gauche, on retrouve le profil de système utilisant de nombreux leviers, où certains individus se distinguent des autres par l'usage de piégeage massif, d'enherbement, d'argile, de paillage au sol et de glu.

Graphique 7 : ACM des leviers - Graphique des individus



En se basant sur l'ACM réalisée sur les leviers qualitatifs, on peut encore une fois distinguer les profils des systèmes. Les systèmes REF se distinguent des ECO et BIO par l'usage des herbicides. Les systèmes ECO et BIO emploient un grand nombre d'autres leviers. Parmi les systèmes ECO et BIO, des distinctions existent dues à l'usage de leviers moins souvent employés (sans qu'ils soient considérés comme des modalités rares). L'usage du piégeage massif, de l'enherbement, et de l'argile caractérisent les systèmes de Sica Centrex (n°18, 20 et 21) et le système BIO du CTIFL (n°9).

En observant les clusters obtenus à partir de cette ACM (Annexe 18, Annexe 19 et Annexe 20), il est possible de distinguer deux groupes de systèmes : un groupe contenant presque tous les systèmes ECO et BIO et caractérisé par l'utilisation de produits de bio contrôle, de suppression des organes touchés, de vigueur et d'absence d'herbicide, et l'autre contenant presque uniquement des systèmes REF et possédant les caractéristiques inverses de celles décrites précédemment. Seul un des systèmes ECO de Sica Centrex se trouve dans le même cluster que les systèmes REF.

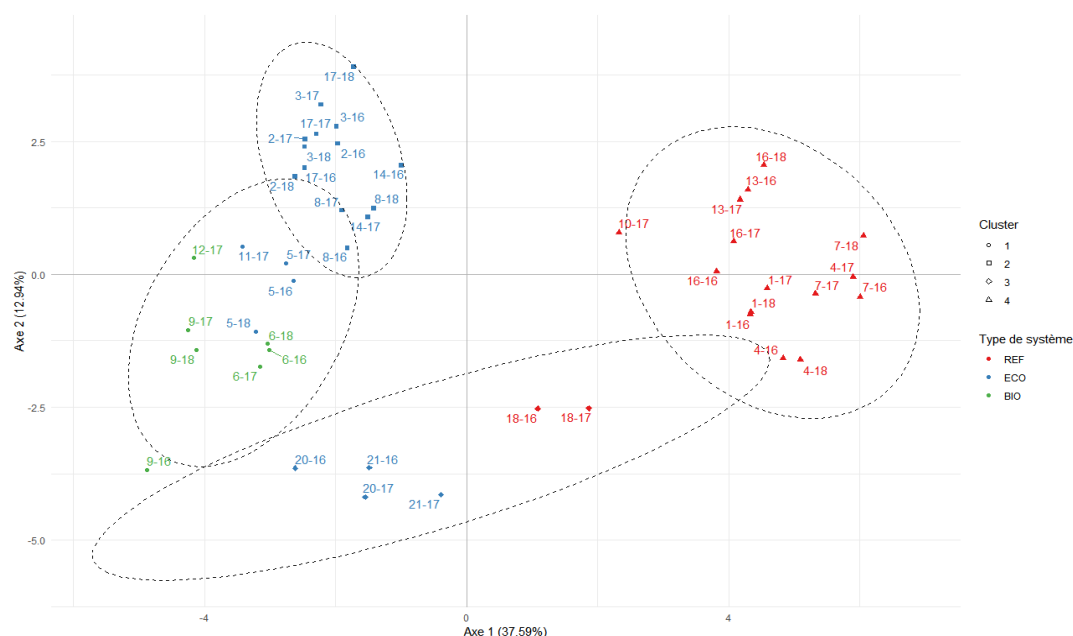
Si lors des analyses précédentes, il était possible de bien discerner les systèmes BIO et REF, ces leviers permettent de distinguer les systèmes REF des systèmes ECO et BIO, sans que ces derniers soient vraiment distincts.

D'autres ACM ont été effectuées, soit avec des critères de performances catégorisés, soit avec ces critères et les leviers qualitatifs, mais les conclusions que nous en avons tiré sont similaires.

6.2.3. Leviers et indicateurs de performance

En se basant sur l’HCPC appliquée sur une AFDM (Annexe 21), il est possible distinguer des groupes d’individus en prenant en compte des variables quantitatives (performances ou explicatives) et qualitatives (leviers).

Graphique 8 : AFDM des leviers et variables agronomiques



Le premier cluster se caractérise par une majorité de système BIO, des leviers tels que le travail au sol et une réduction de la dose, des coûts de production et des IFT verts élevés, mais des rendements, poids de fruits, IFT chimiques ainsi qu’une irrigation et un apport en azote faible. Les individus les plus représentatifs de ce groupe sont les systèmes ECO ainsi que des systèmes BIO. Le deuxième cluster, constitué pour majorité de système ECO, est caractérisé par des IFT herbicides chimiques faibles, l’utilisation de paillage au sol et la gestion de l’irrigation. Il n’y a pas de différence significative de ce groupe par rapport aux performances que nous cherchons à distinguer (rendement, coût, marge partielle, indice réfractométrique). Le troisième cluster se distingue par l’usage de piégeage massif, une pluviométrie, des IFT verts, un indice réfractométrique et un apport en azote élevé, mais une irrigation, et des marges et rendements faibles. Il est à noter que ce cluster contient strictement les systèmes de Sica Centrex. Le quatrième cluster, ne comportant que des systèmes REF, se caractérise par des IFT chimiques (l’usage des herbicides est d’ailleurs le seul levier utilisé y étant significativement associé), une irrigation, des apports en azote et des rendements élevés.

La réalisation d’une ACPVI (Annexe 22) permet également de distinguer des relations entre les performances et les paramètres agronomiques ou économiques des systèmes. Les rendements commercialisables semblent associés positivement aux heures de travail (récolte et hors récolte). L’indice réfractométrique paraît positivement associé à l’apport minéral en azote (et légèrement à la pluviométrie). Le poids moyen des fruits est positivement associé aux IFT chimiques, notamment herbicide, et négativement associé aux IFT verts, dû aux faits que ces différents types d’IFT sont associés à des types de système. Le coût de production est positivement associé aux IFT verts et heures hors récolte. La marge partielle, quant à elle, est associée aux IFT verts, et aux heures (récolte et hors récolte).

6.3. Analyses longitudinales ou groupées

6.3.1. Analyses longitudinales

Pour prendre en compte la dimension temporelle des données, qui avait été mise de côté jusque-là, quelques modèles linéaires mixtes sont réalisés. Les modèles appliqués sur deux variables à expliquer différentes, le rendement commercialisable et la marge partielle, seront retenus (variables les plus synthétiques et paraissant avoir une évolution linéaire). L'aspect de ces variables dans le temps a été vérifié au préalable graphiquement avant d'entreprendre les analyses (Annexe 23 et Annexe 30) et l'adéquation des modèles vérifiée à partir des résidus (Annexe 25, Annexe 27, Annexe 28, Annexe 32, Annexe 34, et Annexe 35).

Pour rappel, le temps t considéré correspond à un temps écoulé depuis la mise en production du verger (première récolte) et seul un intercept aléatoire est ajouté aux modèles. Trois modèles seront considérés successivement :

Le modèle 1 où la variable d'intérêt est uniquement décrite par le temps

$$Y_{ij} = (\beta_0 + b_{0i}) + \beta_1 t_{ij} + \epsilon_{ij}$$

$$\text{Avec } \epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ et } b_i \sim N(0, B)$$

Soit

$$E(Y) = \beta_0 + \beta_1 t$$

Le modèle 2 où l'on ajoutera le type de système comme variable explicative

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 SYST_{2i} + \beta_3 SYST_{3i} + \beta_4 SYST_{2i} t_{ij} + \beta_5 SYST_{3i} t_{ij} + b_{0i} + \epsilon_{ij}$$

$$\text{Avec } \epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ et } b_i \sim N(0, B)$$

Soit

$$E(Y|SYST = REF) = \beta_0 + \beta_1 t$$

$$E(Y|SYST = ECO) = (\beta_0 + \beta_2) + (\beta_1 + \beta_4) t$$

$$E(Y|SYST = BIO) = (\beta_0 + \beta_3) + (\beta_1 + \beta_5) t$$

Et le modèle 3 où le type de système et l'IFT total seront inclus dans le modèle

$$Y_{ij} = \beta_0 + \beta_1 t_{ij} + \beta_2 SYST_{2i} + \beta_3 SYST_{3i} + \beta_4 SYST_{2i} t_{ij} + \beta_5 SYST_{3i} t_{ij}$$

$$+ \beta_6 IFT_{ij} + \beta_7 IFT_{ij} t_{ij} + b_{0i} + \epsilon_{ij}$$

$$\text{Avec } \epsilon_{ij} \sim N(0, \sigma_\epsilon^2) \text{ et } b_i \sim N(0, B)$$

Soit

$$E(Y|SYST = REF) = (\beta_0 + \beta_6 IFT) + (\beta_1 + \beta_7 IFT) t$$

$$E(Y|SYST = ECO) = (\beta_0 + \beta_2 + \beta_6 IFT) + (\beta_1 + \beta_4 + \beta_7 IFT) t$$

$$E(Y|SYST = BIO) = (\beta_0 + \beta_3 + \beta_6 IFT) + (\beta_1 + \beta_5 + \beta_7 IFT)t$$

Après la réalisation de tests du rapport de vraisemblance, les interactions entre le type de système et l'IFT total n'ont pas été jugées significativement différentes de 0 et ont été écartées pour simplifier les modèles ($p = 0.7974$ pour le rendement commercialisable et $p = 0.8639$ pour la marge partielle).

Bien que certaines des variables employées ne présentent parfois pas un intérêt statistique (par exemple, après un test de rapport de vraisemblance, le type de système pourrait être écarté du modèle au profit d'un modèle basé uniquement sur les IFT), ces variables sont conservées car elles importent grandement dans cette étude.

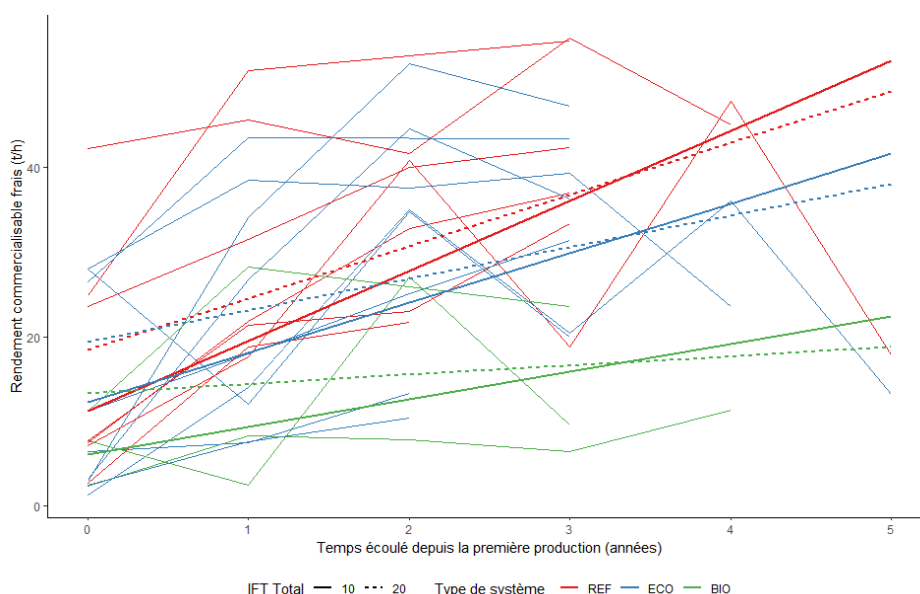
6.3.1.1. *Évolution des rendements commercialisables*

Les estimations des coefficients, effets aléatoires et résultats des tests sont disponibles dans l'Annexe 29. Le rendement augmente significativement dans le temps. Celui-ci est de 17.362 tonnes/ha lors de la première récolte et augmente de 4,462 tonnes/ha à chaque nouvelle récolte (M1)(Annexe 24).

Si l'on considère les types de système (M2), le rendement moyen initial des systèmes REF est supérieur à celui des systèmes ECO, qui est lui-même supérieur à celui des systèmes BIO. La différence des rendements moyens selon les systèmes au temps initial est proche de la significativité ($p=0.068$), notamment de par la différence entre les systèmes REF et BIO. L'évolution moyenne entre les systèmes n'est pas significativement différente de 0 ($p=0.497$)(Annexe 26).

Le rendement moyen est significativement différent seulement au temps initial selon l'IFT total ajusté sur les types des systèmes ($p=0.044$) (M3). Ajusté sur l'IFT total, on note que les systèmes ne présentent pas de différences de rendements significatives.

Graphique 9 : Evolution moyenne des rendements commercialisables (euros/ha) par type des systèmes et IFT prédite par le modèle 3



Malgré le manque de significativité, nous pouvons constater que les systèmes REF et ECO ont des rendements assez proches. Comme le suggère le graphique ci-dessus, l'évolution des rendements

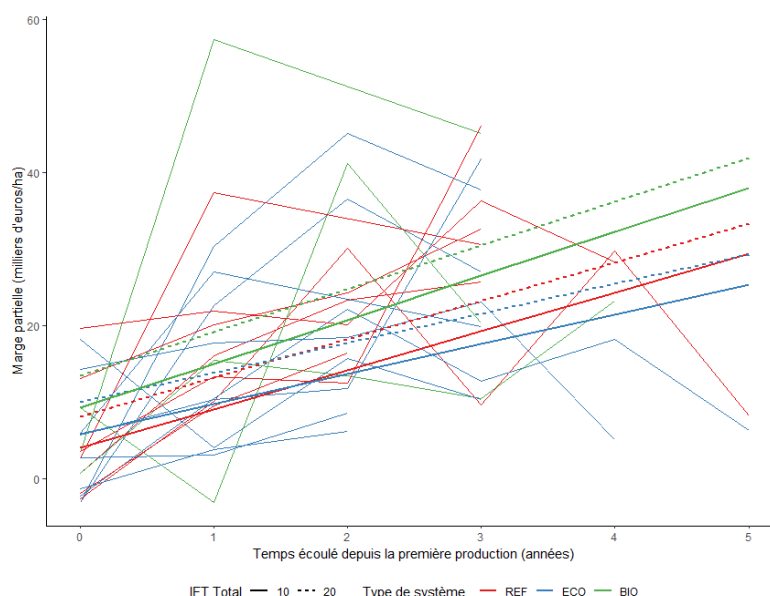
au cours du temps est plus importante pour les systèmes ayant un IFT moins élevé. Si au cours des premières années un IFT élevé assure de meilleurs rendements, ceux-ci évoluent moins, et après un certain temps, les systèmes avec des IFT moins élevés seront plus productifs. Néanmoins, la différence d'évolution imputable aux IFT n'étant pas significativement différente de 0 ($p=0.228$), il n'est pas possible d'affirmer cette constatation et laisse raisonnablement penser que la différence de rendement entre deux doses d'IFT total pour un type de système donné sera constante dans le temps.

6.3.1.2. *Évolution de la marge partielle*

Les estimations des coefficients, effets aléatoires et résultats des tests sont disponibles dans l'Annexe 36.

La marge partielle augmente significativement dans le temps, elle vaut en moyenne 9,5 milliers d'euros à la première récolte et augmente de 4,6 à chaque année supplémentaire (M1)(Annexe 31). En considérant les types de système (M2), les systèmes BIO ont une marge légèrement plus importante que les deux autres systèmes. Néanmoins, il n'y a pas de différence significative entre les systèmes que ce soit à temps initial ($p=0.485$) ou dans le temps ($p=0.739$)(Annexe 33).

Graphique 10: Evolution moyenne de la marge partielle (milliers d'euros/ha) par type des systèmes et IFT prédite par le modèle 3



Sur le modèle comprenant le type de système et l'IFT total (M3), on peut constater graphiquement ci-dessus, que le système BIO a les marges partielles les plus importantes. Les systèmes REF et ECO paraissent très proches pendant les premières années mais la pente des systèmes REF est plus importante. Les IFT n'évoluent quasiment pas dans le temps ($p=0.982$), donc il n'y a que la différence d'IFT initial qui distingue deux systèmes de même type. Malheureusement, aucun des coefficients de ce modèle n'est significativement différent de 0, ce qui ne nous permet pas d'affirmer ces comportements.

7. Discussion

7.1. Réflexions sur les méthodes

7.1.1. Autres méthodes d'analyse envisagées

Certaines méthodes ont été abordées mais n'ont pas été appliquées dans le cadre de ce stage. C'est le cas notamment des PLS, PLS-PM et RGCCA (22,23) et de l'analyse de données groupées.

7.1.1.1. *Partial least squares regression (PLS)*

Bien que régulièrement abordée avec mes maîtres de stage, l'approche PLS n'a finalement pas été retenue. Cette méthode est en effet adaptée pour les données de grande dimension où n le nombre d'individus est inférieur à p le nombre de variables prises en compte dans l'analyse. Or si on ne dispose que de 19 systèmes expérimentaux, le suivi dans le temps fait que l'on dispose non pas de 19, mais de 77 lignes dans la base de données. Sachant que l'on dispose de moins de 80 variables et que certaines ne seront pas prises en compte dans les analyses, on est loin d'un cas de données de grande dimension. En conséquence, les résultats pouvant être issus d'une approche PLS risquent de ne pas être pertinents. Le nombre de variables à prendre en considération aurait d'ailleurs encore diminué car beaucoup sont incomplètes (la PLS ne prenant pas en compte les variables avec des données manquantes). Cette méthode n'était donc pas pertinente dans notre cas.

7.1.1.2. *PLS Path Modeling (PLS-PM) et Regularized Generalized Canonical Correlation Analysis (RGCCA)*

Cependant d'autres méthodes auraient été plus adéquates au vu de la structure de nos données. L'analyse canonique généralisée régularisée (Regularized Generalized Canonical Correlation Analysis ou RGCCA) et l'approche PLS aux modèles à équations structurelles (PLS Path Modeling ou PLS-PM) sont des méthodes proposées pour étudier les relations entre les ensembles de variables observées structurées en un ensemble de blocs. Dans notre cas, on pourrait considérer que chaque bloc est formé par un même ensemble de variables aux différents temps dont on dispose afin de prendre en compte l'aspect pluriannuel des données (qui n'a pu être pris en compte dans les précédentes méthodes mises en application). Si ces méthodes étaient théoriquement adaptées pour les analyses du projet EcoPêche, elles n'ont cependant pas été mises en application durant le stage.

7.1.1.3. *Modèle mixte pour données groupées*

Bien que nos données soient pluriannuelles, elles sont également multisites. Or il n'a pas été possible de prendre en compte à la fois cet effet lors des précédents modèles longitudinaux. En effet, il n'aurait pas été cohérent de le prendre en compte en même temps que le type de système (cela revenait à modéliser la moyenne des différents types de systèmes pour chaque site sachant qu'on ne dispose que de 1 ou 2 types de systèmes par site). Bien que l'on néglige une nouvelle fois l'aspect pluriannuel, il aurait été intéressant de réaliser des analyses de données groupées afin de prendre en compte ce paramètre.

L'emploi de modèle mixte pour données groupées multisites et pluriannuelles est en effet conseillé dans l'ouvrage *De l'analyse des réseaux expérimentaux à la méta-analyse* (24), en prenant en compte des effets aléatoires pour les années, les sites et d'autres variables d'intérêts et les interactions entre celles-ci. Or dans ces modèles, la corrélation intra-système n'est pas présente car, bien que réalisée sur plusieurs années, les systèmes étudiés ne sont pas eux-mêmes suivis dans le temps (culture de plantes annuelles). Cet exemple rappelle que les études pluriannuelles en agronomie basées sur une démarche systémique sont encore peu employées.

7.1.2. Faibles corrélations

Une des remarques que nous pouvons formuler sur l'utilisation des méthodes d'analyse factorielle appliquées ici concerne la corrélation entre les différentes variables utilisées. Une ACP n'a d'intérêt que si les variables sont corrélées entre elles (16). Or, les corrélations entre nos variables sont faibles (comme le présente la matrice des corrélations associée à la première ACP présentée dans ce rapport (Annexe 9), où seules 3 des 28 corrélations sont supérieures à 0.5).

7.1.3. Modèle mixte non linéaire à long terme

Les rendements en fruits, ainsi que l'ensemble des autres performances telles que la marge partielle ont été considérés comme étant des variables linéaires dans le cadre des modèles mixtes. Il faut cependant considérer que ces indicateurs ne sont pas linéaires à long terme. Un verger de pêcheurs a une durée de vie moyenne d'environ 14 ans. Au cours de sa production, il est acquis que ses rendements en fruits augmenteront pendant les premières années de production puis stagneront un certain temps avant de péricliter progressivement (si le verger, alors âgé, n'est pas arraché dès le déclin de sa production). Ainsi, sur le long terme, sa représentation au cours du temps sera plutôt représentative d'une loi normale ayant une faible variance. Si un modèle linéaire mixte peut suffire pour décrire les premières années de production, des modèles plus complexes devront être employés pour des suivis de systèmes plus longs. Il sera important de savoir s'il est possible de maintenir la diminution des intrants sur une plus longue période.

7.1.4. Débat sur la gestion des données manquantes

L'imputation de données manquantes a été une source de débat avec mes collègues de bureau (master ou thésard en agronomie). Jugeant que les valeurs en agronomie dépendent des paramètres climatiques, de l'intensité des bioagresseurs, du comportement des vergers durant les semaines précédentes, il est discutable d'imputer les valeurs manquantes car difficile d'estimer la vraie valeur. Mais, en dehors de renoncer à certaines analyses, les solutions proposées étaient d'écarter les individus incomplets ou de « refaire l'expérience », ce qui, pour un projet réalisé sur plusieurs années, comportant peu d'effectifs et faisant l'objet d'analyses longitudinales, est difficilement envisageable.

7.1.5. Choix des individus

Un système est considéré comme un individu et décrit pour une année donnée par le cumul des performances (souvent rapporté en quantité sur la surface).

Il m'a parfois été conseillé de m'intéresser aux données d'un site en particulier, mais avec une valeur par système-années, il n'est pas possible de faire des analyses (autres que descriptives) sur un même site. Même en prenant les années comme de simples répétitions d'un même système expérimental, il n'est possible de réaliser que des tests non-paramétriques. A l'échelle pluriannuelle ou multisite, les modèles restent limités et peu robustes car construits sur peu d'effectifs. En effet, les modèles construits sont facilement influencés par quelques individus. Pour l'exemple, les quelques systèmes présents à t_5 jouent beaucoup sur les pentes des modèles mixtes (quelques modèles ont été testés séparément sur les quatre et cinq premières années et les pentes obtenues sont des pentes légèrement différentes).

Une solution pour ne pas se retrouver dans ce cas serait de changer d'individu statistique, en choisissant l'arbre comme individu (changement d'échelle). Le suivi par arbre est peut-être plus complexe à mettre en place que le suivi par système, mais au cours du stage j'ai pu noter que des mesures étaient faites sur des arbres ou des rameaux et que ces mesures étaient réalisées sur les mêmes individus à différents temps. Il serait peut-être approprié de récolter des données par arbre en

choisissant aléatoirement un certain nombre d'arbres par système que l'on suivra dans le temps. Ce nombre doit être assez élevé pour pouvoir pallier aux perdus de vue (arbre mort ou coupé pendant le suivi), il faudra aussi écarter les arbres de bordure (souvent influencés par les caractéristiques environnementales) et tenir compte des éventuels gradients d'hétérogénéité.

On disposera d'effectifs plus importants rendant les analyses bien plus solides et évitant les problèmes de convergence des modèles et des analyses plus robustes.

D'autre part, le projet étant prolongé pour 5 années, on disposera d'un suivi plus long pour les systèmes déjà étudiés et quelques systèmes supplémentaires pourraient s'ajouter à ceux déjà suivis. A noter également que certains des systèmes étudiés ne seront plus suivis (arrachage du verger) et que d'autres ont subi des dégâts climatiques au cours de l'année en cours (orage de grêles dans la Drôme) et ne seront pas représentatifs.

7.2. **Réflexions sur les résultats**

D'après les différentes analyses factorielles, les systèmes de référence sont caractérisés par différents IFT chimiques, peu de leviers en dehors de l'utilisation d'herbicide, des heures de travail, des poids de fruits, une irrigation, des apports en azote et des rendements élevés, ainsi que des indices réfractométriques faibles. Les systèmes économes en intrants ont des indices réfractométriques élevés, des IFT herbicides chimiques faibles, beaucoup de leviers de contrôle cultural. Ils se caractérisent par des poids moyens plutôt élevés, mais des rendements moins élevés que les systèmes de référence. Les systèmes biologiques sont caractérisés par des IFT verts, des heures de récoltes et des pourcentages de perte plus élevés. En termes de performance, ses caractéristiques sont des coûts de production élevés et des rendements et poids de fruits faibles.

Les modèles linéaires suggèrent peu de différences entre les systèmes. Les systèmes réduits en intrants paraissent proches des systèmes de référence notamment durant les premières années de production d'un verger. Les rendements et marge partielle des différents systèmes ne sont pas significativement différents à dose d'IFT égale. En termes de rendements, il existe une différence presque significative entre les systèmes, qui n'est cependant plus perceptible après ajustement sur l'IFT total. Pour un type de système donné, les rendements sont parallèles et ne diffèrent que par la dose d'IFT au temps initial (les coefficients du temps et de l'IFT au temps initial sont les seuls à être significativement différents de 0). On peut estimer que le rendement moyen diminue de 0.715 t/ha pour une réduction d'une valeur d'IFT. La marge partielle évolue positivement dans le temps, cependant on ne peut donc pas affirmer ou infirmer d'effet de la réduction d'IFT sur la marge partielle, ajustée sur le type de système.

Il convient de rappeler que la significativité abordée dans ce rapport concerne uniquement l'aspect statistique. Reste à savoir si les différences mises ou non en évidence présentent un aspect significatif en termes d'agronomie.

La recherche en arboriculture présente de nombreuses difficultés. Il existe un ensemble de facteurs environnementaux, climatiques ou, dans notre cas, temporels qui ne sont pas maîtrisables ou impossible à inclure dans les analyses et qui peuvent être des facteurs de confusion importants. De plus, il n'est pas évident de suivre un ensemble strict de consignes que l'on pourrait mettre en place dans le cadre d'un protocole. Certains choix expérimentaux peuvent être modifiés d'une année sur l'autre en fonction des performances obtenues précédemment. Par exemple, un des sites avait fait le choix de réduire de 50% l'usage des intrants, mais devant les faibles performances obtenues, a modifié ses choix en optant pour une réduction d'intrants moins drastique l'année suivante. Il est possible

qu'après avoir étudié les systèmes avec une réduction faible pendant quelques années, d'autres leviers agronomiques ont été introduits, permettant par la suite une réduction plus importante des IFT.

La représentation des leviers n'est pas optimale car, si on sait qu'un levier agronomique a été employé, on ignore l'intensité de son utilisation. Par exemple, si on sait que l'éclaircissage manuel est un des leviers employés, on ignore à quelle fréquence celui-ci est réalisé. Une échelle qualitative plus détaillée aurait été intéressante. La mise en place d'un système de relevé objectif pour que les expérimentateurs puissent retranscrire précisément leurs pratiques chaque année serait une solution envisageable.

Les systèmes ECO de Sica Centrex sont assez différents des autres systèmes ECO, notamment en terme de leviers utilisés et en utilisation d'herbicide. Les différences de ce site tiennent certainement de sa position géographique (seul site situé en Midi-Pyrénées) et des caractéristiques climatique et environnementale qui en découlent. Les règles de décisions appliquées à ce système ne sont donc pas les mêmes. Mais d'autre part, d'après les discussions avec mes encadrants, j'ai pu constater que la gestion des systèmes de ce site avait été assez particulière et avait été entièrement repensée à plusieurs reprises. De ce fait, nous nous sommes interrogés sur l'éventuelle exclusion de ses données.

Le poids des fruits a une importance particulière en termes d'économie car les prix dépendent des calibres obtenus. Le coût de production augmente dès lors que le temps de travail augmente (prix de la main d'œuvre plus couteuse). De ce fait, un système nécessitant plus de contrôle, de manipulations et de techniques en prévention des maladies entraînera forcément des coûts supplémentaires. Les rendements peuvent être associés au temps de travail effectué, mais également à l'irrigation. En effet, la gestion hydrique a des impacts importants sur les systèmes, notamment en termes de qualité des fruits et développement des maladies. Les restrictions hydriques affectent les performances agronomiques (rendement brut, poids moyen des fruits et calibres) des arbres, tandis qu'à l'opposé, les déficits hydriques faibles à modérés réduisent la proportion de fruits malades. La forte réduction de l'irrigation a été jugée bénéfique pour les vergers ne faisant pas de protection par les fongicides (25). Une réduction à l'approche de la récolte peut être favorable à la qualité, mais peut diminuer le calibre (26). L'apport en azote semble avoir de l'importance sur l'indice réfractométrique. La marge est plus importante pour les systèmes BIO du fait que cette appellation permet des prix de vente plus élevés que les autres types de systèmes.

On ne dispose pas ici d'informations sur le sol ou la qualité des terrains. Les pêcheurs sont pourtant très sensibles au type de sol sur lequel ils sont plantés. Il est avéré que les pêcheurs sont sensibles à la fatigue du sol, à l'asphyxie, aux herbicides résiduels (utilisés pour désherber les rangs) et que le rendement est corrélé positivement avec la profondeur de sol exploité (26). On considérera que le type de porte-greffe employé est adapté aux terrains et limite les pertes de performances liées au sol.

Il n'a pas été possible de prendre en compte la variété, bien que les rendements et la résistance aux bioagresseurs (et donc les traitements nécessaires) y soient fortement liés.

Des discussions m'ont permis de constater que les systèmes de référence mis en place dans le projet EcoPêche (au moins en ce qui concerne celui de l'INRA d'Avignon) se distinguent des systèmes conventionnels car l'usage de produits phytosanitaires a déjà tendance à être réduit par rapport à ces derniers. Tenant compte de cela, il est possible que les systèmes de référence utilisés ne soient pas représentatifs des systèmes étudiés. Il faut aussi reconsidérer les variables économiques dont on dispose. Les sites expérimentaux n'étant pas autorisés à écouler les productions sur le marché, les marges sont ici calculées d'après un ensemble de prix théoriques selon les calibres des fruits et ce pour

les différents sites et années. Il est possible que ces valeurs soit légèrement distinctes des prix du marché.

Il est couramment admis que le faible coût des pesticides tend à favoriser leur utilisation, de même que le temps de travail nécessaire à leur application. L'article *Reducing pesticide use while preserving crop productivity and profitability on arable farms* (27) montre que la réduction des pesticides est possible sans perte de rentabilité (les pesticides peuvent être réduits de 42% sans effet négatif sur la productivité et la rentabilité dans 59% des fermes arables). Bien que cet article ne soit pas articulé autour des vergers de pêchers, il est possible d'en retenir une des conclusions. A savoir, que la réduction de pesticides sans coûts supplémentaires est possible mais aussi sans intérêts financiers supplémentaires (pas de réduction des coûts) et que la transition vers un usage réduit des pesticides nécessite la mise en place de stratégies qui augmentent la complexité de management des cultures.

Conclusion

Conclusion du stage

L'objectif de ce stage était de proposer et de mettre en œuvre une démarche statistique adaptée afin d'exploiter au mieux les données du projet EcoPêche et d'étudier les performances de systèmes économes en intrants par rapport à des systèmes de culture conventionnelle en vergers de pêchers.

Mais parfois, la façon dont sont menées les expériences, la multiplicité des facteurs à prendre en compte qu'ils soient explicatifs ou à expliquer, facteurs d'information ou de confusion, rendent la démarche statistique bien plus complexe que l'on peut s'y attendre initialement.

Il n'a pas été possible de distinguer des performances différentes entre les systèmes étudiés. A dose d'IFT égale, les systèmes suivent des trajectoires parallèles ou non-significativement différentes.

Si ces résultats vont dans le sens des attentes du projet et du plan Ecophyto, on peut émettre des réserves sur ces résultats. La prise en compte des aspects temporels et multisites sur des essais systémiques, le choix de considérer les systèmes comme individus statistiques, l'absence de répétitions et le faible effectif ont complexifié les analyses qui auraient été adaptées à ce contexte. La réalisation d'analyses factorielles n'a pas permis de prendre en compte des corrélations liées aux systèmes, aux sites et temps, tandis que les modèles mixtes linéaires étaient restreints en raison d'un effectif faible, de suivi court (peu d'observation par sujet) et à l'absence de répétitions par sites. La poursuite du projet en prenant en compte l'échelle de l'arbre et non du système permettrait des analyses plus robustes.

L'étude d'essai système à la fois multicentrique et pluriannuel étant peu utilisé actuellement, ce stage aura néanmoins permis d'aborder les différentes méthodes d'analyses statistiques pouvant être utilisées dans ce contexte.

Conclusion personnelle

Ce stage a été très instructif et m'a permis d'avoir une approche plus poussée de la recherche. J'ai pu en apprendre davantage sur la recherche agronomique. C'était pour moi un autre aspect de la recherche, directement lié à la santé publique en raison des pratiques actuelles, sur lequel je souhaitais me pencher davantage depuis quelques années.

J'ai pu quelquefois me rendre sur les vergers et participer à la récolte d'informations, ce qui donnait un aspect plus concret à mes missions car j'en visualisais l'ampleur et le travail fourni.

La recherche en agronomie m'a paru plus complexe car la démarche mise en place ne suit pas la même logique. En santé publique, le protocole rédigé avant le lancement d'une étude va aborder les objectifs, les informations à prendre en compte et les méthodes d'analyses prévues. Ce cadre préalablement défini ne m'a pas paru présent dans la recherche agronomique. Certes, il existe des protocoles expérimentaux rédigés en amont, mais ceux-ci sont centrés sur les manipulations et expérimentations à réaliser et n'abordent pas précisément les aspects statistiques à venir. En conséquence, après la récolte des données, il faut rechercher les méthodes adaptées. Cela peut s'avérer complexe car les données n'ont pas forcément été récoltées de manière à faciliter ou rendre possibles les méthodes les plus adéquates au contexte. J'ai souvent eu l'impression d'une dualité entre les aspects agronomiques et les aspects statistiques.

Au début de mon stage, j'ai été assez déroutée car j'imaginais que les essais systémiques suivaient un fonctionnement proche des essais factoriels que l'on avait pu aborder dans mes enseignements, que les systèmes et pratiques agricoles seraient les mêmes sur l'ensemble des sites

(seuls les facteurs climatiques auraient différencié les systèmes de différents sites) et qu'il y aurait des répétitions de systèmes sur un même site. La mise en place du projet et le choix de l'échelle des données m'ont assez déstabilisée. Cette confrontation entre la théorie vue au cours de ma formation et la réalité des pratiques était assez imposante et m'a posé quelques problèmes d'adaptation.

Je compte poursuivre ma carrière dans la recherche, mais probablement plus dans la phase de préparation des études, notamment sur la mise en place et la gestion de bases de données et les aspects de data management. J'ai pu constater que si la plupart de mes collègues de bureau avaient de grandes capacités en agronomie (de par leur formation), les bases de la création de base de données (méthode Merise) étaient loin d'être acquises et posent problème pour la suite des analyses. Ces aspects sont essentiels dans la recherche et je compte bien les mettre en pratique à l'avenir.

Bibliographie

1. INRA. Rapport d'activité 2017. 2018 Jun.
2. Ministère de la culture et de l'alimentation. Qu'est-ce que l'agro-écologie ? . Disponible en ligne sur : <https://agriculture.gouv.fr>
3. INRA. Des partenariats ancrés au cœur des territoires. 2017. Disponible en ligne sur : <http://institut.inra.fr>
4. Association Française De Protection Des Plantes. Répertoire terminologique en protection des plantes. 2017.
5. Portail de la protection intégrée, EcophytoPIC. Les bioagresseurs | EcophytoPIC . Disponible en ligne sur : <http://www.ecophytopic.fr/arbo/surveillance/les-bioagresseurs>
6. Baldi I., Cordier S., Coumoul X., et al. Pesticides: Effets sur la santé (expertise scientifique collective). INSERM; 2013 Jul.
7. Ministère de la Transition écologique et solidaire. Directive n° 91/414/CEE du 15/07/91 concernant la mise sur le marché des produits phytopharmaceutiques (Texte abrogé) . Disponible en ligne sur : <https://aida.ineris.fr>
8. Le Roux X., Barbault R., Baudry J., et al. Agriculture et biodiversité : Valoriser les synergies (expertise scientifique collective). INRA; 2008 Jun.
9. Ministère de l'agriculture, de l'agroalimentaire et de la forêt, Ministère de l'écologie, du développement durable et de l'énergie. Plan Ecophyto II. 2015 Oct.
10. Cretin L., Triquenot A. Apports de produits phytopharmaceutiques en arboriculture : nombre de traitements et indicateur de fréquence de traitements. Agreste; 2018 Aug. (Les Dossiers). Report No.: 43.
11. Lichou J., Mandrin J.F., Breniaux D. Production fruitière intégrée des fruits à noyau. CTIFL; 2001.
12. Expertise scientifique collective I– C. Pesticides, agriculture et environnement. Editions Quæ; 2011.
13. Plénet D., Hilaire C., Blanc P., et al. EcoPêche : Conception et évaluation multisite de vergers de pêche –nectarine économes en produits phytopharmaceutiques et en intrants. 2019;76.
14. Ministère de l'agriculture et de l'alimentation. Agreste - La statistique, l'évaluation et la prospective agricole - Glossaire. Disponible en ligne sur : <http://agreste.agriculture.gouv.fr/definitions/glossaire>
15. Ministère de l'Agriculture et de l'Alimentation. Infos rapides – Fruits – Pêche. Agreste Infos rapides. 2018 Sep;(129).
16. Husson F., Lê S., Pagès J. Analyse de données avec R. 2e éd. revue et augmentée. Rennes: Presses universitaires de Rennes; 2016. (Pratique de la statistique).
17. Husson F. R pour la statistique et la science des données. Rennes: Presses universitaires de Rennes; 2018. (Pratique de la statistique).

18. Husson F, Josse J, Pages J. Principal component methods - hierarchical clustering - partitional clustering: why would we need to choose for visualizing data? 2010;
19. Pagès J. Analyse factorielle de données mixtes. *Revue de statistique appliquée*. 2004;52(4):93–111.
20. Van den Wollenberg AL. Redundancy analysis an alternative for canonical correlation analysis. *Psychometrika*. 1977 Jun;42(2):207–19.
21. Jacquet K, Prodon R. Analyses multivariées avec ade4 dans R. :32.
22. Tenenhaus A, Tenenhaus M. Analyse canonique généralisée régularisée et approche PLS. 2010;
23. Tenenhaus M, Gauchi J.-P., Ménardo C. Régression PLS et applications. *Revue de statistique appliquée*. 1995;43(1).
24. Makowski D., Piraux F., Brun F. De l'analyse des réseaux expérimentaux à la méta-analyse. Méthodes et applications avec le logiciel R pour les sciences agronomiques et environnementales. 2018. (Savoir faire).
25. Plénet D., Simon S., Vercambre G., et al. Systèmes de culture en arboriculture fruitière et qualité des fruits. 2010;10.
26. Guiheneuf Y. Productions fruitières. 1998.
27. Lechenet, M., Dessaint, F, Py, G., et al. Reducing pesticide use while preserving crop productivity and profitability on arable farms. *Nature Plants*. 2017;3.

Annexes

| | |
|---|----|
| Annexe 1 : Liste des systèmes | 43 |
| Annexe 2 : Disposition géographique des sites du projet EcoPêche (Carte réalisée sous R)..... | 44 |
| Annexe 3 : Relations entre l'ensemble des variables liées à l'IFT présentes dans la base de données | 44 |
| Annexe 4 : Tableau des caractéristiques générales des systèmes EcoPêche | 45 |
| Annexe 5 : Différence annuelle d'IFT total entre les systèmes par rapport au système de référence. | 46 |
| Annexe 6 : Comparaison des leviers agronomiques (moyenne, écart-type) utilisés sur les différents types de systèmes selon les années et les types de leviers | 47 |
| Annexe 7 : Comparaisons des coûts de production entre les systèmes de référence et les autres systèmes selon les différents sites | 48 |
| Annexe 8 : Comparaisons des marges partielles entre les systèmes de référence et les autres systèmes selon les différents sites | 48 |
| Annexe 9 : Matrice des corrélations, critère de Kaiser et de Cattell, graphiques des variables et des individus obtenus par défaut sous R pour les ACP des IFT, heures de travail et poids moyens des fruits | 48 |
| Annexe 10 : HCPC basé sur l'ACP des IFT, heures de travail et poids moyens des fruits– Sortie R par défaut | 51 |
| Annexe 11 : Graphiques des clusters basés sur l'ACP des IFT, heures de travail et poids moyens des fruits | 51 |
| Annexe 12 : Code et sortie R - Caractéristiques des clusters basés sur l'ACP des IFT, heures de travail et poids moyen des fruits | 52 |
| Annexe 13 : ACP portant sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits..... | 53 |
| Annexe 14 : HCPC basé sur l'ACP portant sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits – Sortie R par défaut | 54 |
| Annexe 15 : Graphiques des clusters basés sur l'ACP sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits..... | 55 |
| Annexe 16 : Code et sortie R - Caractéristiques des clusters basés sur l'ACP sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits | 55 |
| Annexe 17 : ACM des leviers - graphique des variables..... | 56 |
| Annexe 18 : HCPC basé sur l'ACM des leviers –Sortie R par défaut..... | 57 |
| Annexe 19 : ACM des leviers - Graphique des individus avec clusters avec type de systèmes | 58 |
| Annexe 20 : Code et sortie R - Caractéristiques des clusters basés sur l'ACM des leviers | 58 |
| Annexe 21 : AFDM des leviers et variables agronomiques – Graphiques de l'ensemble des variables, des variables quantitatives, des modalités, des individus, sortie de l'HCPC et caractéristiques des clusters | 59 |
| Annexe 22 : ACPVI des performances et des variables explicatives | 65 |
| Annexe 23 : Evolution des rendements commercialisables (euros/ha) globale et par type de système | 66 |
| Annexe 24 : Evolution moyenne des rendements commercialisables (euros/ha) prédits par le modèle 1 | 67 |
| Annexe 25 : Analyse des résidus - Rendements commercialisables - Modèle 1..... | 67 |

| | |
|---|----|
| Annexe 26 : Evolution moyenne des rendements commercialisables (euros/ha) par type des systèmes prédite par le modèle 2..... | 68 |
| Annexe 27 : Analyse des résidus - Rendements commercialisables - Modèle 2..... | 68 |
| Annexe 28 : Analyse des résidus - Rendements commercialisables - Modèle 3..... | 69 |
| Annexe 29 : Estimation des effets fixes, des effets aléatoires et résultats du test du rapport de vraisemblance sur les modèles - Modèle mixte sur le rendement commercialisable..... | 70 |
| Annexe 30 : Evolution de la marge partielle (euros/ha) globale et par type de système..... | 71 |
| Annexe 31 : Evolution moyenne de la marge partielle (milliers d'euros/ha) prédit par le modèle 1 .. | 72 |
| Annexe 32 : Analyse des résidus - Marge partielle- Modèle 1..... | 72 |
| Annexe 33: Evolution moyenne de la marge partielle (milliers d'euros/ha) par type des systèmes prédite par le modèle 2..... | 73 |
| Annexe 34 : Analyse des résidus - Marge partielle - Modèle 2..... | 73 |
| Annexe 35 : Analyse des résidus - Marge partielle - Modèle 3..... | 74 |
| Annexe 36 : Estimation des effets fixes, des effets aléatoires et résultats du test du rapport de vraisemblance sur les modèles - Modèle mixte sur la marge partielle..... | 75 |

| Identifiant du système | Site | Type de système |
|------------------------|----------------|-----------------|
| 1 | INRA Avignon 1 | REF |
| 2 | INRA Avignon 1 | ECO (ECO1) |
| 3 | INRA Avignon 1 | ECO (ECO2) |
| 4 | INRA Gotheron | REF |
| 5 | INRA Gotheron | ECO |
| 6 | INRA Gotheron | BIO |
| 7 | CTIFL | REF |
| 8 | CTIFL | ECO |
| 9 | CTIFL | BIO |
| 10 | SEFRA | REF |
| 11 | SEFRA | ECO |
| 12 | SEFRA | BIO |
| 13 | SERFEL 1 | REF |
| 14 | SERFEL 1 | ECO |
| 16 | SERFEL 2 | REF |
| 17 | SERFEL 2 | ECO |
| 18 | Sica CENTREX | REF (Enh) |
| 20 | Sica CENTREX | ECO (Enh) |
| 21 | Sica CENTREX | ECO (Desh) |

Annexe 1 : Liste des systèmes

Annexe 2 : Disposition géographique des sites du projet EcoPêche (Carte réalisée sous R)



Annexe 3 : Relations entre l'ensemble des variables liées à l'IFT présentes dans la base de données

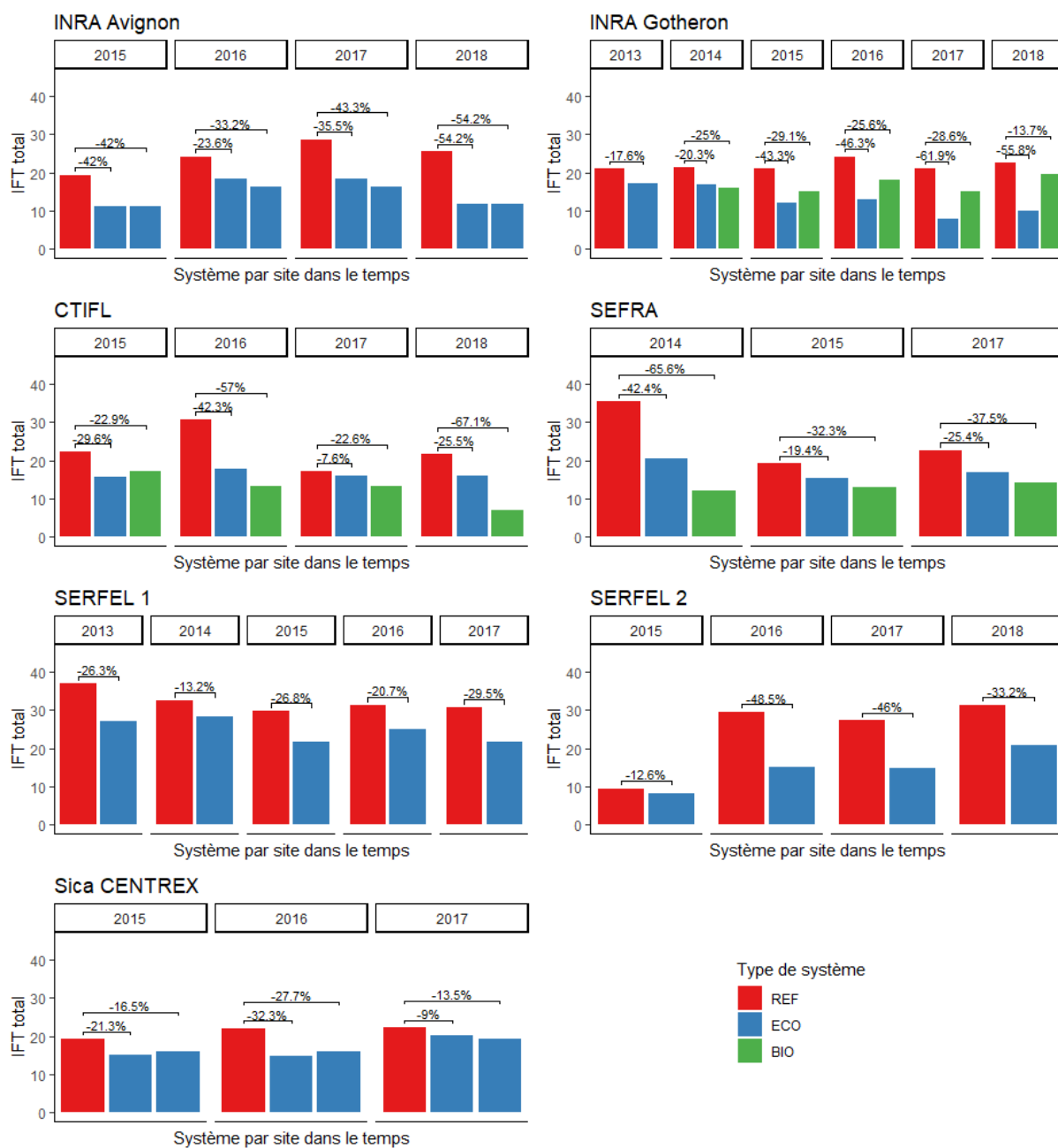
$$\begin{aligned}
 \text{IFT total} &= \text{IFT herbicides total} + \text{IFT fongicides total} + \text{IFT insecticides total} \\
 &\quad + \text{IFT autres produits total} \\
 &= \text{IFT herbicides chimiques} \\
 &\quad + (\text{IFT fongicides chimiques} + \text{IFT fongicides vert}) \\
 &\quad + (\text{IFT insecticides chimiques} + \text{IFT insecticides verts}) \\
 &\quad + (\text{IFT autres produits chimiques} + \text{IFT autres produits verts}) \\
 &= \text{IFT chimiques total} + \text{IFT verts total}
 \end{aligned}$$

| | | Global | Sites | | | | | | | Type de système | | |
|---------------------------|-------------------|-----------|-----------------|------------------|-----------|-----------|-----------|-----------|--------------|-----------------|-----------|----------|
| | | | INRA Avignon | INRA Gotheron | CTIFL | SEFRA | SERFEL 1 | SERFEL 2 | Sica Centrex | REF | ECO | BIO |
| N | | 19 | 3 | 3 | 3 | 3 | 2 | 2 | 3 | 7 | 9 | 3 |
| Variété (%) | Elise | 2 (10.5) | 0 (0.0) | 2 (66.7) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (11.1) | 1 (33.3) |
| | Nectardream | 3 (15.8) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 1 (33.3) |
| | Nectarlove | 3 (15.8) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 2 (22.2) | 0 (0.0) |
| | Orine | 3 (15.8) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 1 (14.3) | 2 (22.2) | 0 (0.0) |
| | Sandine | 2 (10.5) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 0 (0.0) |
| | Surprise | 1 (5.3) | 0 (0.0) | 1 (33.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 0 (0.0) | 0 (0.0) |
| | Sweet Star | 3 (15.8) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 1 (33.3) |
| | Western Red | 2 (10.5) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 0 (0.0) |
| | Porte-greffe (%) | Cadaman | 6 (31.6) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 2 (28.6) | 3 (33.3) |
| | GF 677 | 9 (47.4) | 3 (100.0) | 0 (0.0) | 3 (100.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (42.9) | 4 (44.4) | 2 (66.7) |
| | Montclar | 4 (21.1) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 2 (100.0) | 0 (0.0) | 2 (28.6) | 2 (22.2) | 0 (0.0) |
| Type de fruits (%) | Nectarine Blanche | 8 (42.1) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 2 (100.0) | 0 (0.0) | 3 (42.9) | 4 (44.4) | 1 (33.3) |
| | Nectarine Jaune | 5 (26.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 0 (0.0) | 3 (100.0) | 2 (28.6) | 3 (33.3) | 0 (0.0) |
| | Pêche Blanche | 6 (31.6) | 0 (0.0) | 3 (100.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (28.6) | 2 (22.2) | 2 (66.7) |
| Groupe de précocité (%) | Saison | 6 (31.6) | 3 (100.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (28.6) | 3 (33.3) | 1 (33.3) |
| | Tardive | 13 (68.4) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 3 (100.0) | 2 (100.0) | 2 (100.0) | 3 (100.0) | 5 (71.4) | 6 (66.7) | 2 (66.7) |
| Année de plantation (%) | 2011 | 5 (26.3) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 0 (0.0) | 0 (0.0) | 2 (28.6) | 2 (22.2) | 1 (33.3) |
| | 2012 | 3 (15.8) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 1 (33.3) |
| | 2013 | 11 (57.9) | 3 (100.0) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 3 (100.0) | 4 (57.1) | 6 (66.7) | 1 (33.3) |
| Densité de plantation (%) | 476 | 2 (10.5) | 0 (0.0) | 0 (0.0) | 1 (33.3) | 1 (33.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (28.6) | 0 (0.0) | 0 (0.0) |
| | 533 | 3 (15.8) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 1 (33.3) |
| | 556 | 5 (26.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (66.7) | 2 (100.0) | 1 (50.0) | 0 (0.0) | 2 (28.6) | 2 (22.2) | 1 (33.3) |
| | 571 | 5 (26.3) | 2 (66.7) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 2 (28.6) | 3 (33.3) | 0 (0.0) |
| | 889 | 1 (5.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (50.0) | 0 (0.0) | 0 (0.0) | 1 (11.1) | 0 (0.0) |
| | 909 | 1 (5.3) | 1 (33.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (11.1) | 0 (0.0) |
| | 1010 | 2 (10.5) | 0 (0.0) | 0 (0.0) | 2 (66.7) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (11.1) | 1 (33.3) |
| | 2Y | 9 (47.4) | 2 (66.7) | 0 (0.0) | 1 (33.3) | 3 (100.0) | 2 (100.0) | 1 (50.0) | 0 (0.0) | 5 (71.4) | 3 (33.3) | 1 (33.3) |
| | | Gobelet | 6 (31.6) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 3 (100.0) | 2 (28.6) | 3 (33.3) |
| Forme fruitière (%) | Palmette | 1 (5.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (50.0) | 0 (0.0) | 0 (0.0) | 1 (11.1) | 0 (0.0) |
| | Y oblique | 3 (15.8) | 1 (33.3) | 0 (0.0) | 2 (66.7) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 2 (22.2) | 1 (33.3) |
| | 2013 | 4 (21.1) | 0 (0.0) | 2 (66.7) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 0 (0.0) | 0 (0.0) | 2 (28.6) | 2 (22.2) | 0 (0.0) |
| Année de campagne (%) | 2014 | 4 (21.1) | 0 (0.0) | 1 (33.3) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (14.3) | 1 (11.1) | 2 (66.7) |
| | 2015 | 11 (57.9) | 3 (100.0) | 0 (0.0) | 3 (100.0) | 0 (0.0) | 0 (0.0) | 2 (100.0) | 3 (100.0) | 4 (57.1) | 6 (66.7) | 1 (33.3) |
| Age du verger (%) | 3 | 18 (94.7) | 3 (100.0) | 2 (66.7) | 3 (100.0) | 3 (100.0) | 2 (100.0) | 2 (100.0) | 3 (100.0) | 7 (100.0) | 9 (100.0) | 2 (66.7) |
| | 4 | 1 (5.3) | 0 (0.0) | 1 (33.3) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 0 (0.0) | 1 (33.3) |

Annexe 4 : Tableau des caractéristiques générales des systèmes EcoPêche

Cette description ne comporte pas de tests car le peu d'effectif oblige à réaliser des tests de Fisher (2x2) pour la majorité des variables.

Annexe 5 : Différence annuelle d'IFT total entre les systèmes par rapport au système de référence

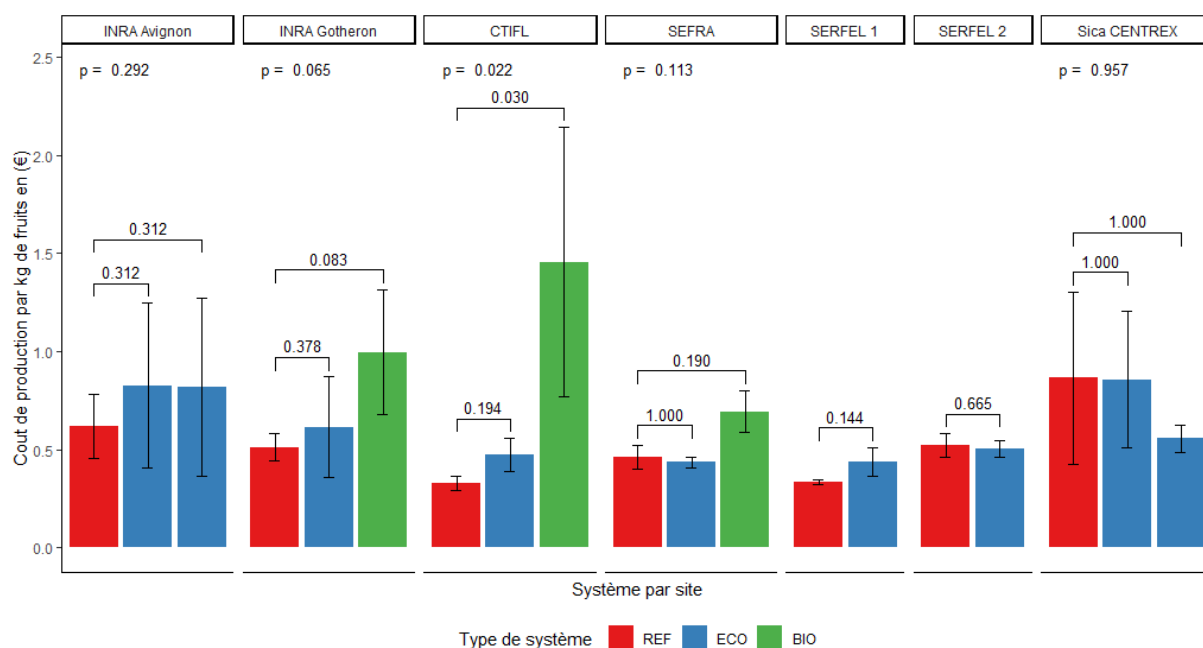


| | Années de suivi | | | | | | | | | | | | | | | | | | | | | | | |
|-------------------|-----------------|--------|------|--------|------|--------|--------|------|--------|------|--------|------|--------|--------|------|--------|------|--------|------|--------|--------|--|--|--|
| | 2016 | | | | | | | | 2017 | | | | | | | | 2018 | | | | | | | |
| | REF | | ECO | | BIO | | p* | REF | | ECO | | BIO | | p* | REF | | ECO | | BIO | | p* | | | |
| | | | | | | | | | | | | | | | | | | | | | | | | |
| N | 6 | | 8 | | 2 | | | 7 | | 9 | | 3 | | | 4 | | 5 | | 2 | | | | | |
| Biologique | 0,67 | (0,52) | 2,00 | (0,00) | 2,00 | (0,00) | <0,001 | 0,86 | (0,69) | 1,89 | (0,33) | 2,00 | (0,00) | 0,001 | 0,50 | (0,58) | 2,00 | (0,00) | 2,00 | (0,00) | <0,001 | | | |
| Biotechnique | 0,17 | (0,41) | 0,25 | (0,46) | 0,50 | (0,71) | 0,690 | 0,14 | (0,38) | 0,22 | (0,44) | 0,33 | (0,58) | 0,818 | 0,00 | (0,00) | 0,00 | (0,00) | 0,50 | (0,71) | 0,092 | | | |
| Chimique | 0,83 | (0,41) | 0,12 | (0,35) | 0,00 | (0,00) | 0,006 | 1,00 | (0,00) | 0,22 | (0,44) | 0,00 | (0,00) | <0,001 | 1,00 | (0,00) | 0,00 | (0,00) | 0,00 | (0,00) | <0,001 | | | |
| Contrôle cultural | 2,50 | (1,52) | 9,00 | (1,20) | 8,50 | (0,71) | <0,001 | 2,71 | (1,70) | 8,89 | (1,27) | 8,67 | (0,58) | <0,001 | 1,75 | (0,96) | 9,60 | (0,89) | 8,50 | (0,71) | <0,001 | | | |
| Physique | 0,33 | (0,52) | 1,12 | (0,35) | 2,00 | (1,41) | 0,007 | 0,29 | (0,49) | 1,33 | (0,71) | 2,33 | (1,15) | 0,002 | 0,25 | (0,50) | 1,00 | (0,00) | 2,00 | (1,41) | 0,025 | | | |
| Raisonnement | 0,00 | (0,00) | 1,12 | (0,64) | 1,50 | (0,71) | 0,002 | 0,14 | (0,38) | 1,11 | (0,78) | 1,67 | (0,58) | 0,005 | 0,00 | (0,00) | 1,40 | (0,55) | 1,50 | (0,71) | 0,004 | | | |

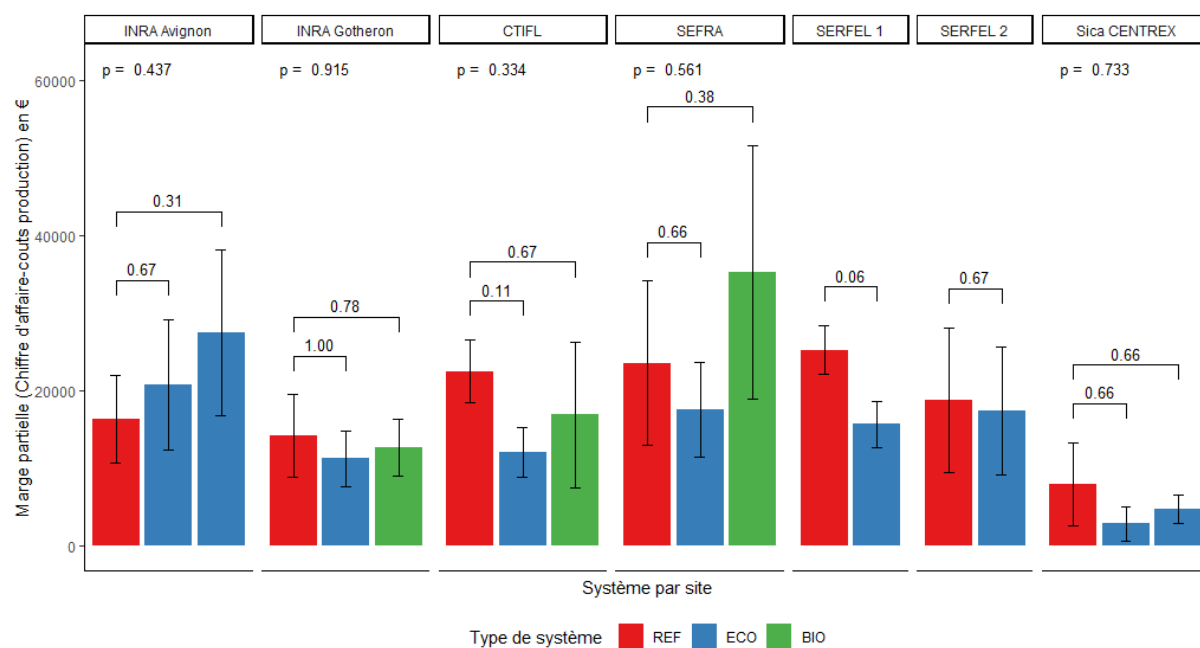
*Les p-values sont basé sur le test de Kruskal-Wallis. Les mêmes tests ont été effectué pour comparer les types de systèmes deux à deux.

Annexe 6 : Comparaison des leviers agronomiques (moyenne, écart-type) utilisés sur les différents types de systèmes selon les années et les types de leviers

Annexe 7 : Comparaisons des coûts de production entre les systèmes de référence et les autres systèmes selon les différents sites



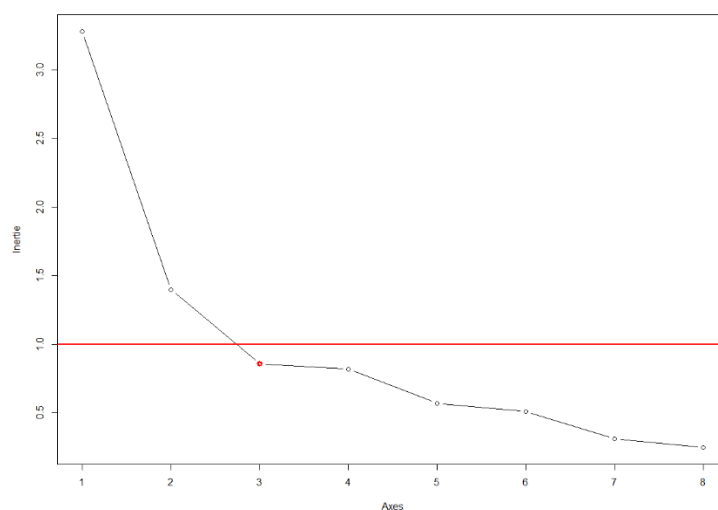
Annexe 8 : Comparaisons des marges partielles entre les systèmes de référence et les autres systèmes selon les différents sites



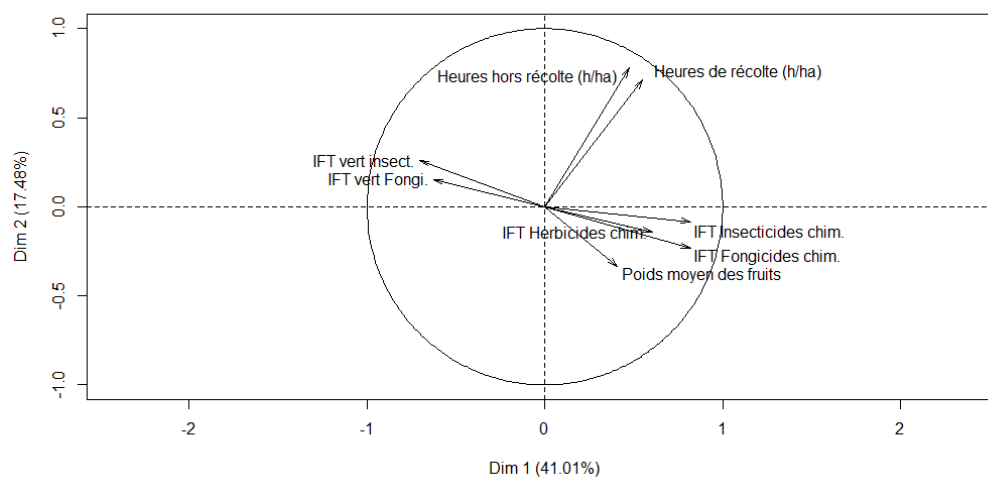
Annexe 9 : Matrice des corrélations, critère de Kaiser et de Cattell, graphiques des variables et des individus obtenus par défaut sous R pour les ACP des IFT, heures de travail et poids moyens des fruits

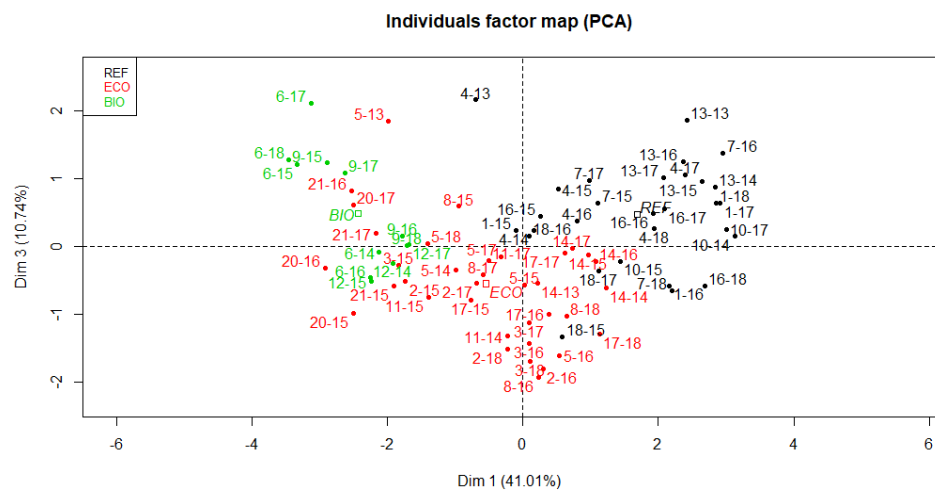
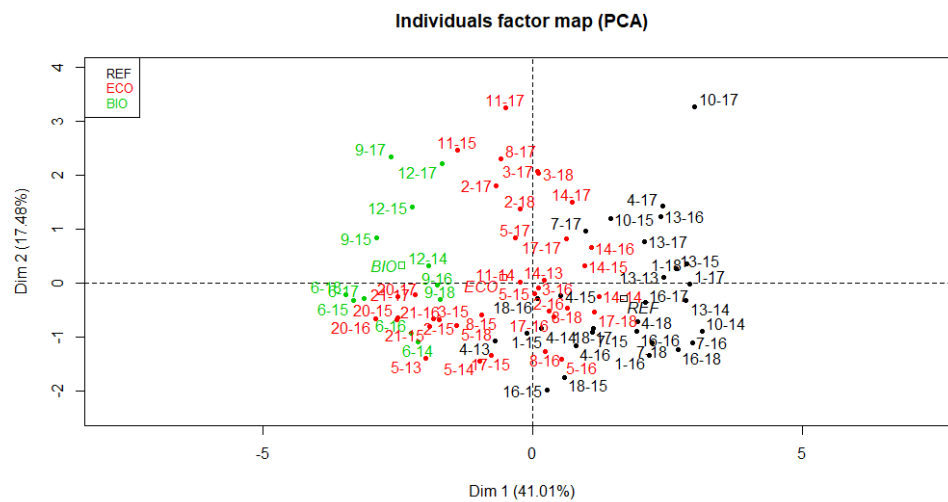
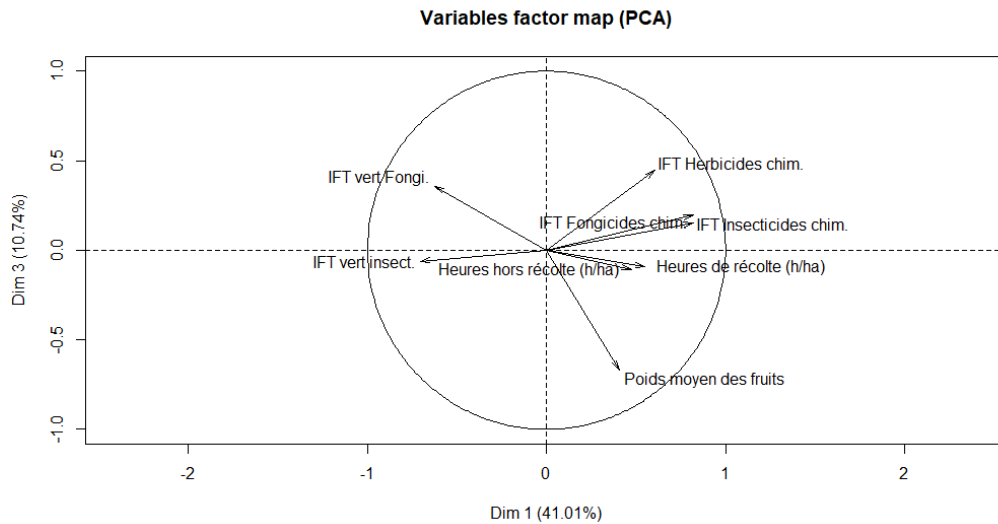
| | IFT Herbicides chim. | IFT Fongicides chim. | IFT Insecticides chim. | IFT vert Fongi. | IFT vert insect. | Heures récolte (h/ha) | de | Heures hors récolte (h/ha) | Poids moyen des fruits |
|----------------------|----------------------------|----------------------------|------------------------------|-----------------------|------------------------|-----------------------------|----|-------------------------------------|---------------------------|
| IFT Herbicides chim. | 1 | 0.4409427 | 0.4722863 | -0.1999807 | -0.33583872 | 0.16490171 | | 0.19802334 | 0.20280043 |

| | | | | | | | | |
|----------------------------|------------|------------|------------|------------|-------------|-------------|-------------|-------------|
| IFT Fongicides chim. | 0.4409427 | 1.0000000 | 0.6935072 | -0.4762442 | -0.55723751 | 0.27449592 | 0.17750299 | 0.25931696 |
| IFT Insecticides chim. | 0.4722863 | 0.6935072 | 1.0000000 | -0.3752749 | -0.49499387 | 0.31196404 | 0.31673772 | 0.26469789 |
| IFT vert Fongi. | -0.1999807 | -0.4762442 | -0.3752749 | 1.0000000 | 0.41535067 | -0.23266989 | -0.17604599 | -0.25459879 |
| IFT vert insect. | -0.3358387 | -0.5572375 | -0.4949939 | 0.4153507 | 1.00000000 | -0.27986935 | -0.07444545 | -0.23696740 |
| Heures de récolte (h/ha) | 0.1649017 | 0.2744959 | 0.3119640 | -0.2326699 | -0.27986935 | 1.00000000 | 0.66685929 | 0.03964251 |
| Heures hors récolte (h/ha) | 0.1980233 | 0.1775030 | 0.3167377 | -0.1760460 | -0.07444545 | 0.66685929 | 1.00000000 | 0.06611064 |
| Poids moyen des fruits | 0.2028004 | 0.2593170 | 0.2646979 | -0.2545988 | -0.23696740 | 0.03964251 | 0.06611064 | 1.00000000 |

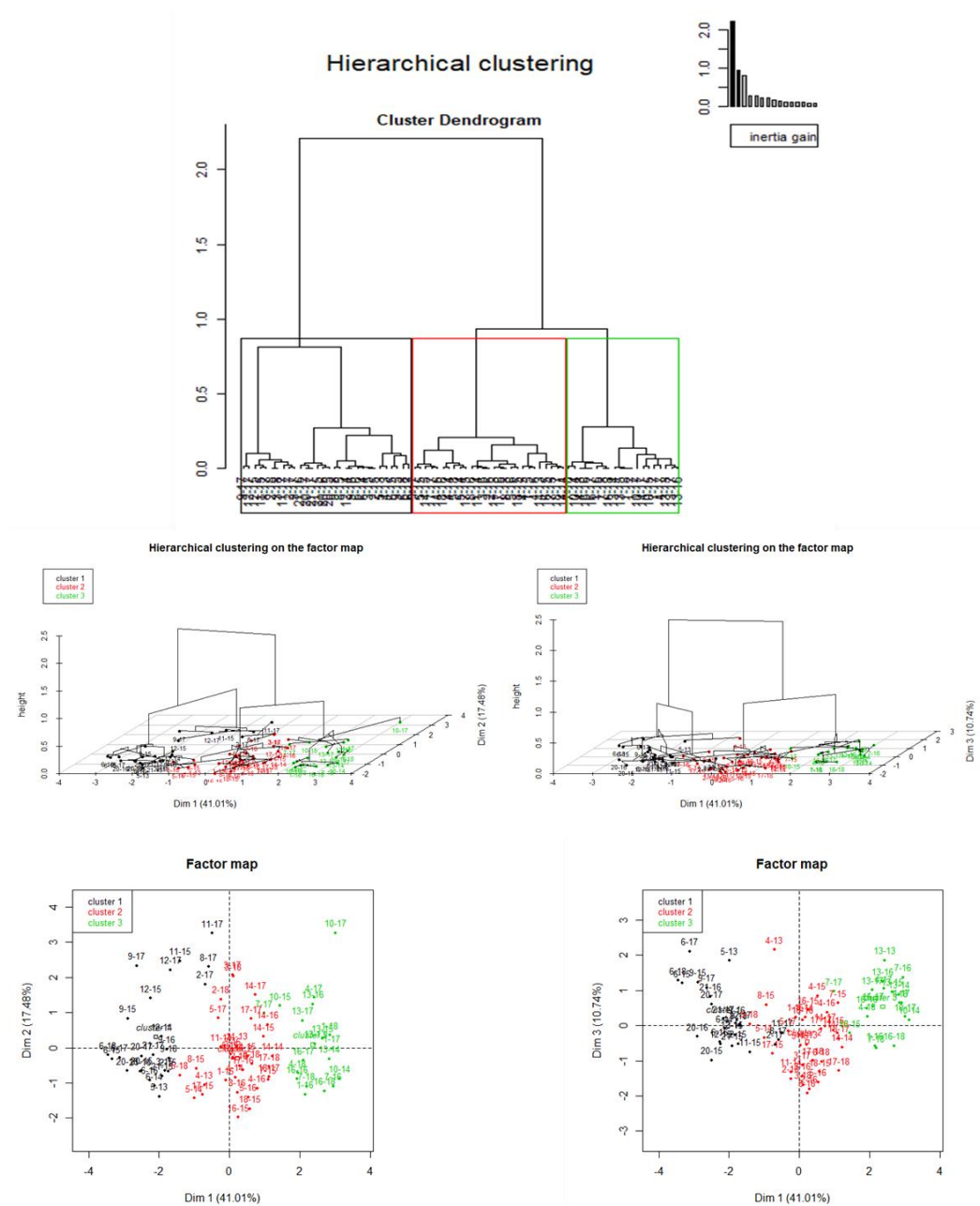


Variables factor map (PCA)

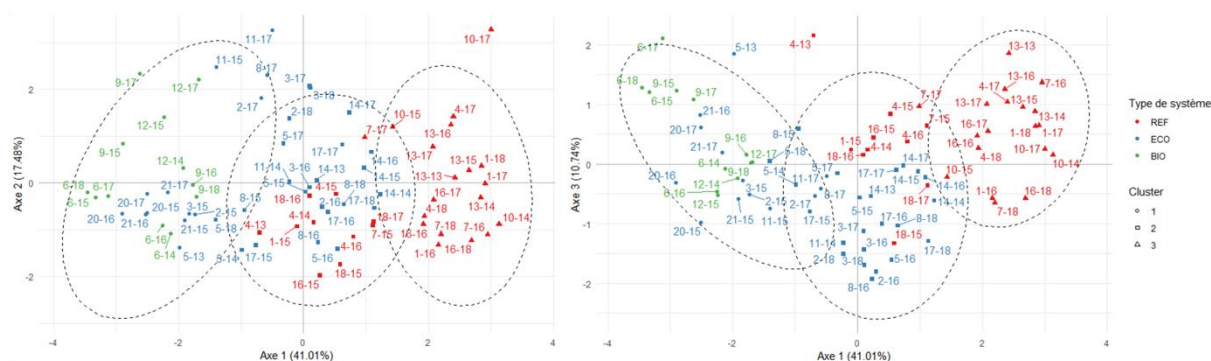




Annexe 10 : HCPC basé sur l'ACP des IFT, heures de travail et poids moyens des fruits– Sortie R par défaut



Annexe 11 : Graphiques des clusters basés sur l'ACP des IFT, heures de travail et poids moyens des fruits



Annexe 12 : Code et sortie R - Caractéristiques des clusters basés sur l'ACP des IFT, heures de travail et poids moyen des fruits

```
># Description par les variables
>hcpc$desc.var$quanti
$`1`
v.testMean in categoryOverallmeansd in categoryOverallstdp.value
IFT.vert.insect.      6.244823      5.643867      3.3601299      1.9887049      2.2105537      4.242812e-10
IFT.vert.Fongi.       5.942337      4.385067      2.5900866      1.3115670      1.8259021      2.809868e-09
IFT.Herbicides.chim.  -3.066999      0.163600      0.4969968      0.3473083      0.6570871      2.162199e-03
Poids.moyen.des.fruits -3.156759      159.197982      170.8203133      21.5929136      22.2549670      1.595331e-03
Heures.de.récolte.(h/ha) -3.343550      215.280433      317.7465780      185.8245335      185.2455183      8.271380e-04
IFT.Insecticides.chim. -4.874970      0.782400      4.2190476      1.2342100      4.2612611      1.088248e-06
IFT.Fongicides.chim.  -6.240866      4.235008      8.5939163      1.8478427      4.2219009      4.351546e-10

$`2`
v.testMean in categoryOverallmeansd in categoryOverallstdp.value
IFT.vert.insect.      -2.566793      2.6085859      3.3601299      1.4078551      2.2105537      0.0102643983
IFT.Herbicides.chim.  -2.712613      0.2609091      0.4969968      0.5078139      0.6570871      0.0066755059
IFT.vert.Fongi.       -3.558837      1.7293939      2.5900866      1.4354008      1.8259021      0.0003725013

$`3`
v.testMean in categoryOverallmeansd in categoryOverallstdp.value
IFT.Herbicides.chim.      6.444878      1.345724      0.4969968      0.4240006      0.6570871      1.156938e-10
IFT.Insecticides.chim.    6.314079      9.611404      4.2190476      3.8111364      4.2612611      2.717755e-10
IFT.Fongicides.chim.      5.005028      12.828834      8.5939163      3.1107144      4.2219009      5.585388e-07
Heures.hors.récolte.(h/ha) 3.557864      521.881579      391.4493478      151.3332322      182.9217007      3.738834e-04
Heures.de.récolte.(h/ha)  3.191851      436.247206      317.7465780      143.7421742      185.2455183      1.413645e-03
IFT.vert.Fongi.          -2.369057      1.723158      2.5900866      1.1796389      1.8259021      1.783351e-02
IFT.vert.insect.         -3.836332      1.660526      3.3601299      0.6518353      2.2105537      1.248857e-04

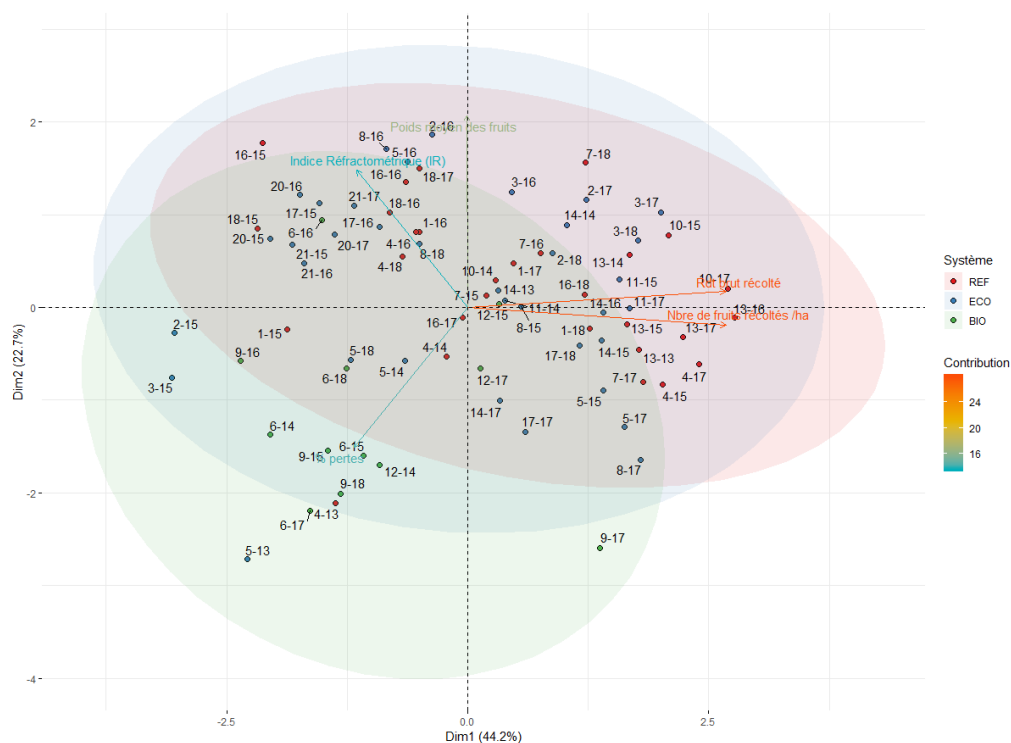
>hcpc$desc.var$category
$`1`
Cla/ModMod/Cla Global p.valuev.test
Type.de.système=BIO 100 48 15.58442 1.415075e-07 5.263280
Type.de.système=REF 0 0 37.66234 2.667760e-07 -5.145523

$`2`
Cla/ModMod/Cla Global p.valuev.test
Type.de.système=ECO 63.88889 69.69697 46.75325 0.0005715352 3.444775
Type.de.système=BIO 0.00000 0.00000 15.58442 0.0005739074 -3.443655

$`3`
Cla/ModMod/Cla Global p.valuev.test
Type.de.système=REF 65.51724 100 37.66234 3.944858e-11 6.606134
Type.de.système=BIO 0.00000 0 15.58442 2.426700e-02 -2.252875
Type.de.système=ECO 0.00000 0 46.75325 4.818567e-07 -5.033400

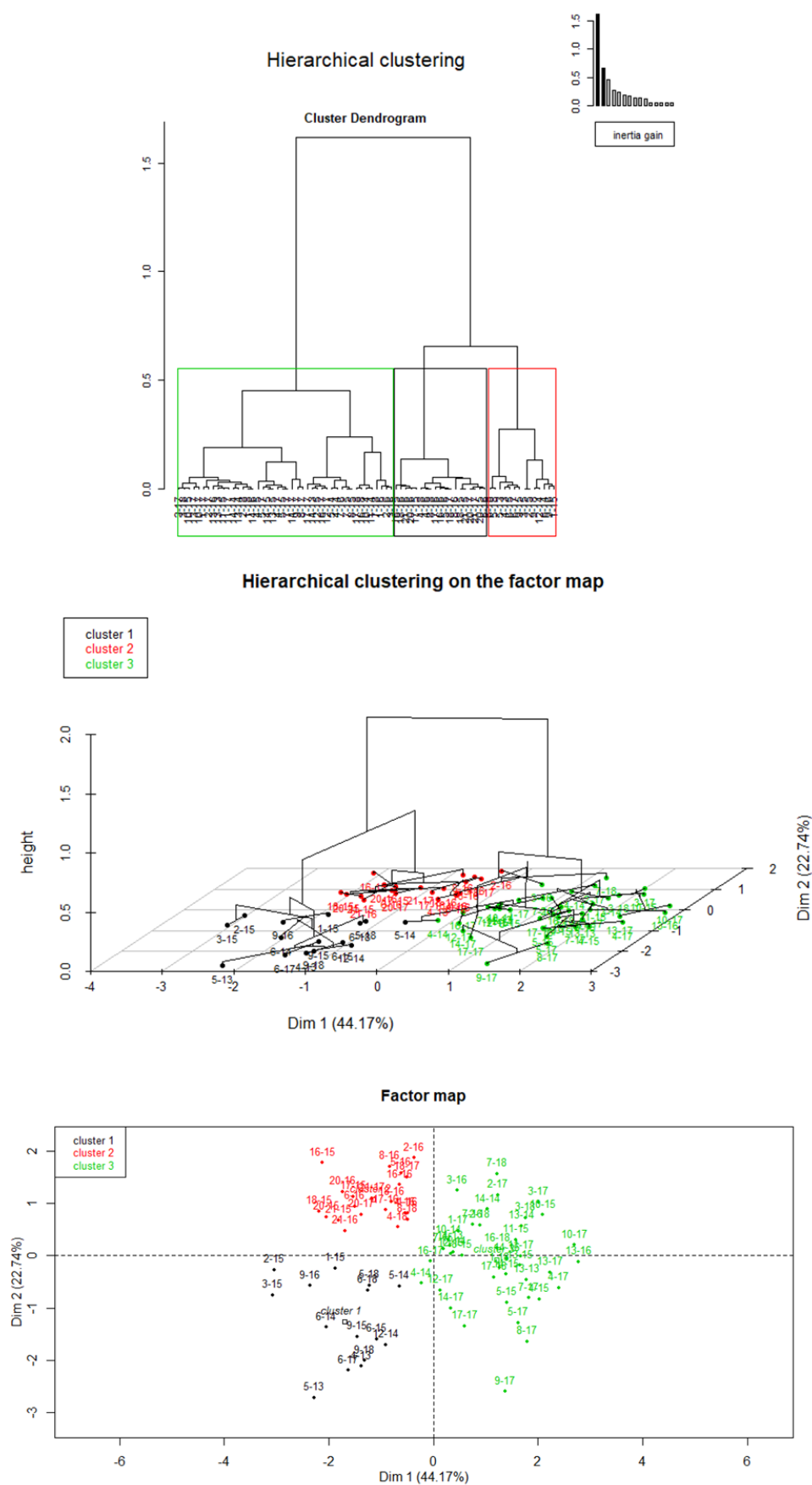
>
># Parangons des clusters
>hcpc$desc.ind$para
Cluster: 1
 9-16 12-14 3-15 12-15 2-15
0.7549321 0.8363119 1.3264176 1.4242235 1.4599473
-----
Cluster: 2
18-16 17-16 8-18 4-14 14-14
0.8112778 0.8350131 0.9637911 1.0303183 1.1519584
-----
Cluster: 3
16-17 13-14 4-18 16-16 1-18
0.6053737 0.9696060 1.0497993 1.1390918 1.1986664
```

Annexe 13 : ACP portant sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits

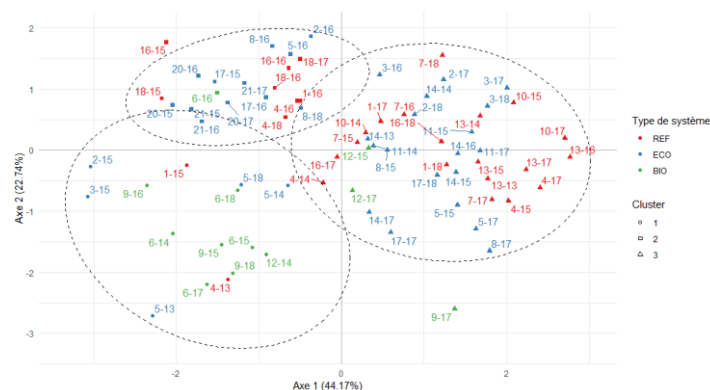


L'axe 1, qui explique 44,7 % de l'inertie, est fortement corrélé aux rendements bruts et au nombre de fruits/ha. Le poids moyen des fruits est corrélé avec l'axe 2, tandis que l'indice réfractométrique et le pourcentage de perte sont liés aux 2 axes, mais la représentativité de ces 3 variables n'est pas optimale.

Annexe 14 : HCPC basé sur l'ACP portant sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits – Sortie R par défaut



Annexe 15 : Graphiques des clusters basés sur l'ACP sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits



Annexe 16 : Code et sortie R - Caractéristiques des clusters basés sur l'ACP sur le rendement brut, le nombre de fruits/ha, l'indice réfractométrique, le pourcentage de perte et le poids moyen des fruits

```
># Description par les variables
>hpcpc$desc.var$quanti
$`1`
v.testMean in categoryOverallmeansd in categoryOverallstdp.value
%.pertes          5.468360      35.12329      18.7618      13.291333      12.82987  4.542181e-08
Poids.moyen.des.fruits -3.682362     151.70872     170.8203     24.039119     22.25497  2.310834e-04
Nb.de.fruits.récoltés./ha -4.274307     72075.14702    174823.3791    35169.028199  103077.82446  1.917328e-05
Rdt.brut.récolté      -4.655666      11.11106      29.6337      5.822784      17.05995  3.229355e-06

$`2`
v.testMean in categoryOverallmeansd in categoryOverallstdp.value
Indice.Réfractométrique.(IR) 4.227640      13.30591      11.9608      1.201053  1.698574e+00  2.361554e-05
Poids.moyen.des.fruits      3.582111     185.75319     170.8203     14.510133  2.225497e+01  3.408294e-04
%.pertes                    -2.050937     13.83287      18.7618      7.995991  1.282987e+01  4.027310e-02
Rdt.brut.récolté            -4.073063     16.61773      29.6337      8.891475  1.705995e+01  4.639882e-05
Nb.de.fruits.récoltés./ha    -4.456019     88785.46151    174823.3791    44452.543838  1.030778e+05  8.349546e-06

$`3`
v.testMean in categoryOverallmeansd in categoryOverallstdp.value
Nb.de.fruits.récoltés./ha    7.370321     256482.39748    174823.3791    64556.918545  1.030778e+05  1.702181e-13
Rdt.brut.récolté            7.331205      43.07699      29.6337     10.033814  1.705995e+01  2.280926e-13
%.pertes                    -2.509975     15.30046      18.7618      9.454919  1.282987e+01  1.207397e-02
Indice.Réfractométrique.(IR) -3.680746     11.28879      11.9608      1.507781  1.698574e+00  2.325527e-04

>hpcpc$desc.var$category
$`1`
Cla/ModMod/Cla  Global  p.valuev.test
Type.de.système=BIO 66.666667 53.33333 15.58442 0.000108297 3.871209
Type.de.système=REF 6.896552 13.33333 37.66234 0.030751733 -2.160271

$`2`
NULL

$`3`
Cla/ModMod/Cla  Global  p.valuev.test
Type.de.système=BIO 25 7.317073 15.58442 0.04006074 -2.053122

>
># Description par les individus
>hpcpc$desc.ind$para
Cluster: 1
      6-18      9-16      5-18      4-13      6-14
```

0.9154673 0.9808401 1.4661499 1.4957072 1.5669857

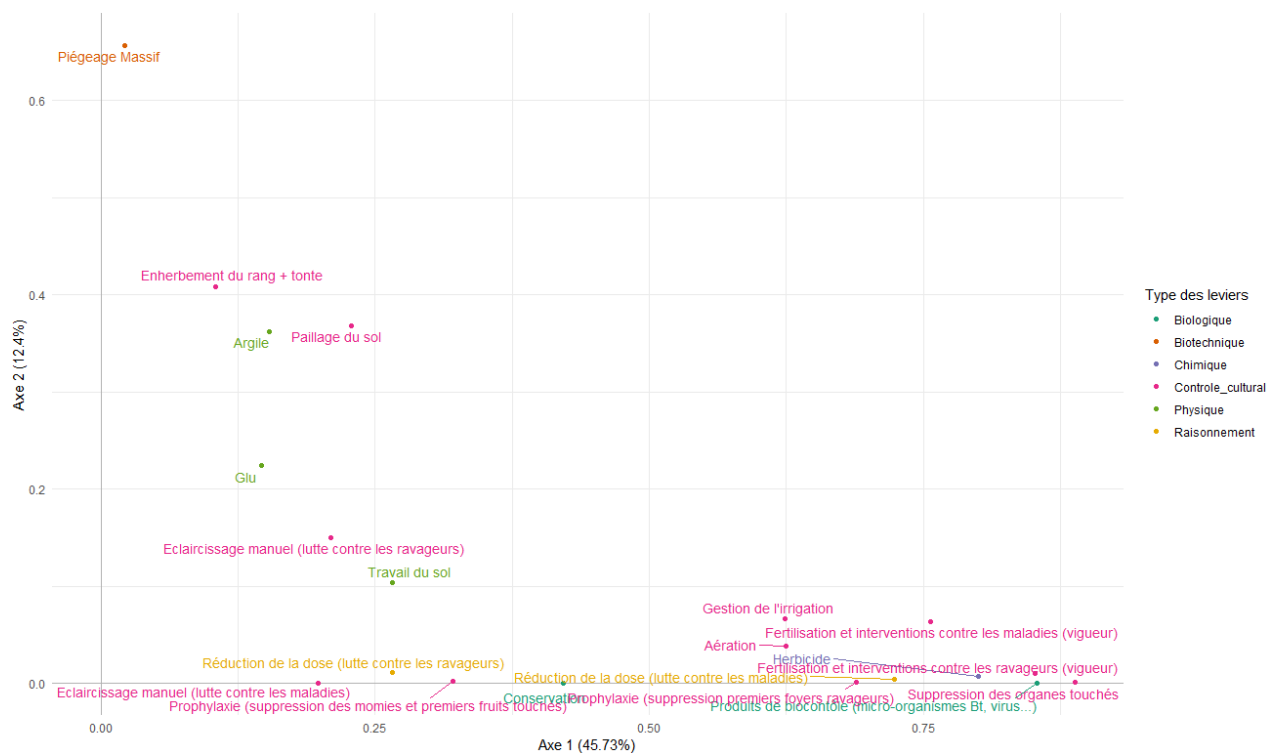
Cluster: 2

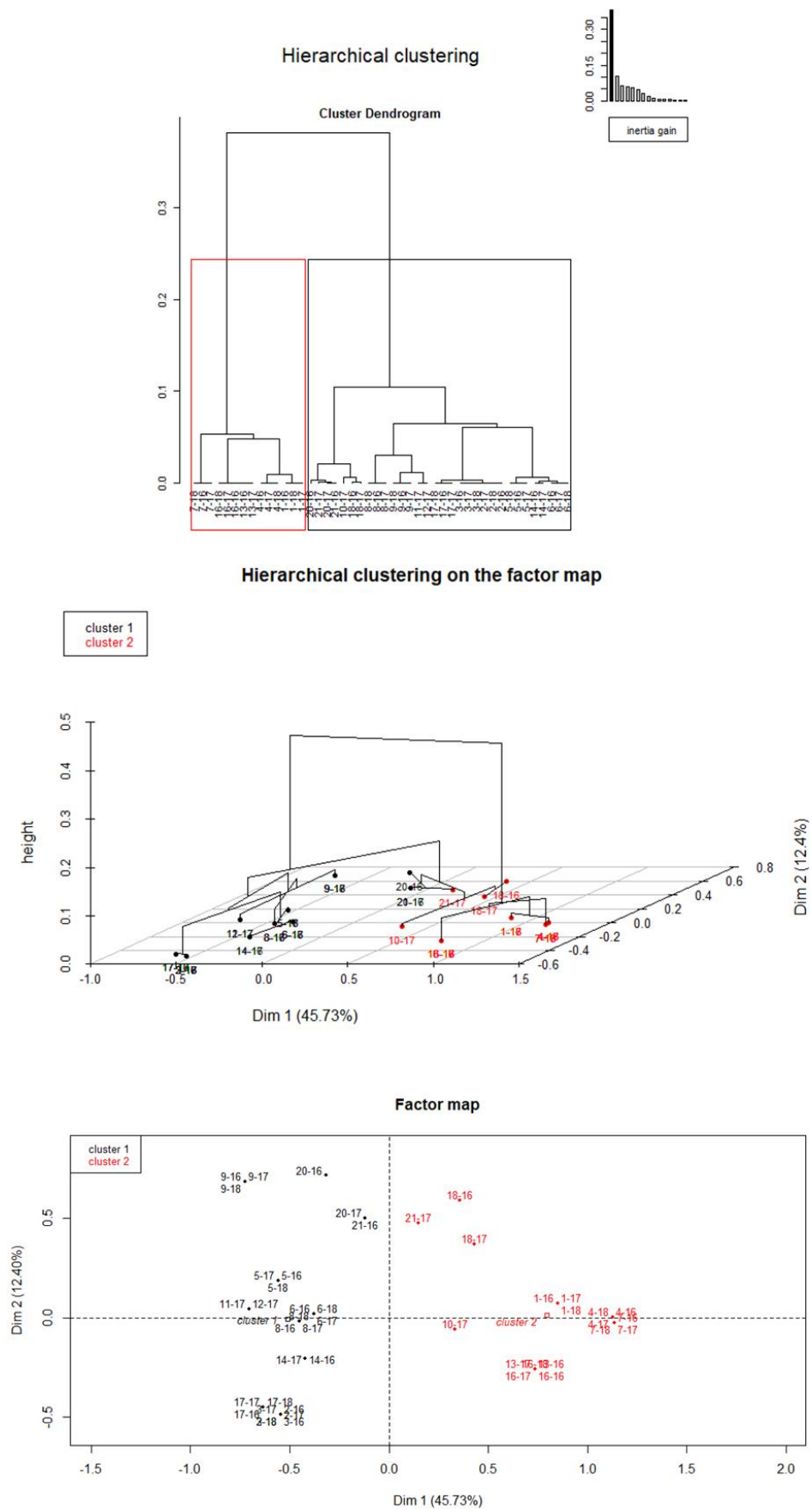
21-17 20-17 17-16 5-16 6-16
0.4182574 0.5970287 0.7214102 0.7830094 0.7928344

Cluster: 3

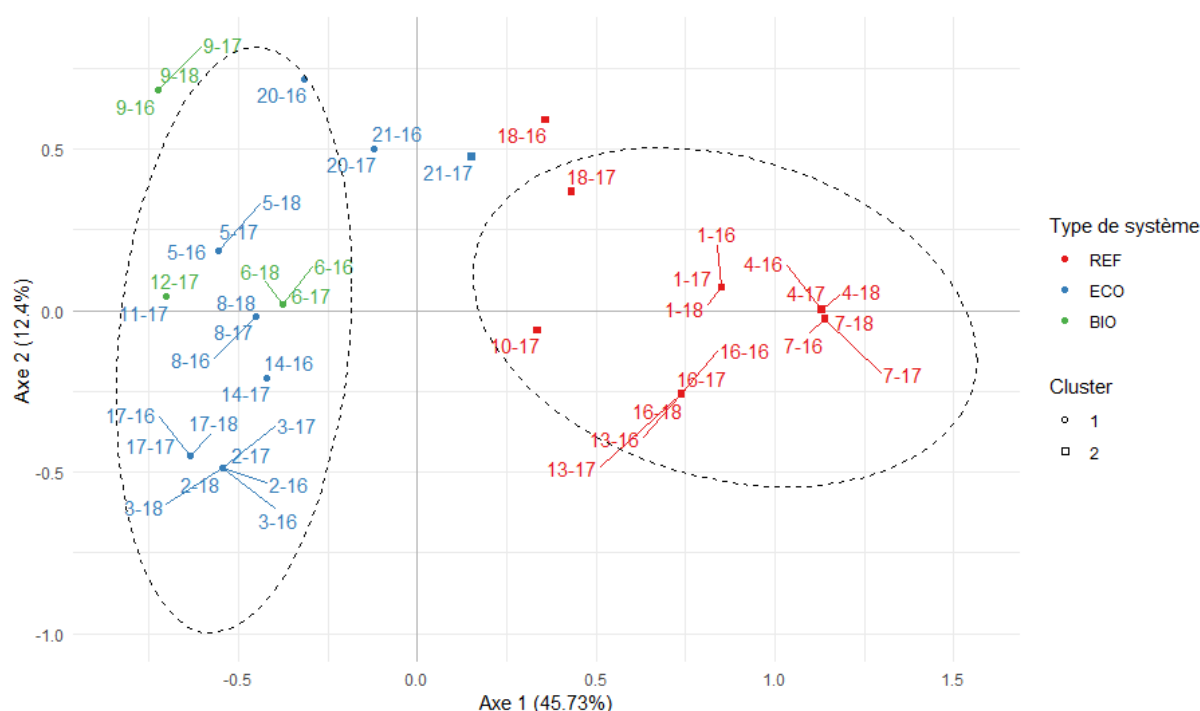
1-18 14-16 5-15 11-15 7-17
0.4574157 0.7395079 0.8710654 0.8711259 0.9500741

Annexe 17 : ACM des leviers - graphique des variables





Annexe 19 : ACM des leviers - Graphique des individus avec clusters avec type de systèmes



Annexe 20 : Code et sortie R - Caractéristiques des clusters basés sur l'ACM des leviers

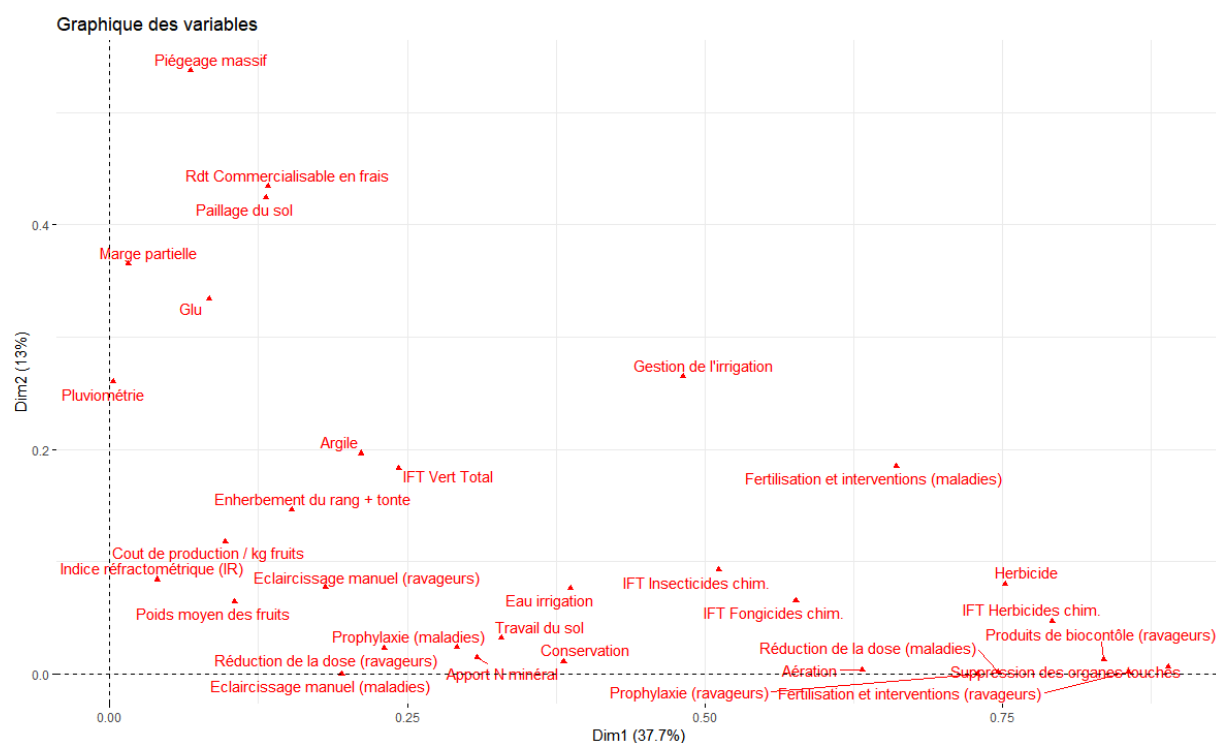
```
> hcpc$desc.var$category$`1`[c(1:10,(nrow(hcpc$desc.var$category$`1`)-10):(nrow(hcpc$desc.var$category$`1`))),]
Cla/ModMod/Cla Global p.valuev.test
r_produits_biocontrôle=r_produits_biocontrôle_1 96.55172 100.000000 63.04348 1.028751e-11 6.802422
m_suppression_organes_touche=s_m_suppression_organes_touche_1 96.55172 100.000000 63.04348 1.028751e-11 6.802422
r_vigueur=r_vigueur_1 90.32258 100.000000 67.39130 1.594564e-09 6.034499
m_reduction_dose=m_reduction_dose_1 96.29630 92.857143 58.69565 1.651322e-09 6.028848
a_herbicide=a_herbicide_0 96.29630 92.857143 58.69565 1.651322e-09 6.028848
r_prophylaxie=r_prophylaxie_1 96.15385 89.285714 56.52174 1.064934e-08 5.720049
m_vigueur=m_vigueur_1 96.15385 89.285714 56.52174 1.064934e-08 5.720049
m_gestion_irrigation=m_gestion_irrigation_1 95.65217 78.571429 50.00000 8.475058e-07 4.924097
m_aeration=m_aeration_1 80.00000 100.000000 76.08696 2.385467e-06 4.717682
syst_req=ECO 95.45455 75.000000 47.82609 2.796598e-06 4.685220
a_travail_sol=a_travail_sol_0 45.45455 53.571429 71.73913 3.679232e-04 -3.562084
m_aeration=m_aeration_0 0.00000 0.000000 23.91304 2.385467e-06 -4.717682
m_gestion_irrigation=m_gestion_irrigation_0 26.08696 21.428571 50.00000 8.475058e-07 -4.924097
r_prophylaxie=r_prophylaxie_0 15.00000 10.714286 43.47826 1.064934e-08 -5.720049
m_vigueur=m_vigueur_0 15.00000 10.714286 43.47826 1.064934e-08 -5.720049
m_reduction_dose=m_reduction_dose_0 10.52632 7.142857 41.30435 1.651322e-09 -6.028848
a_herbicide=a_herbicide_1 10.52632 7.142857 41.30435 1.651322e-09 -6.028848
r_vigueur=r_vigueur_0 0.00000 0.000000 32.60870 1.594564e-09 -6.034499
syst_req=REF 0.00000 0.000000 36.95652 1.028751e-11 -6.802422
r_produits_biocontrôle=r_produits_biocontrôle_0 0.00000 0.000000 36.95652 1.028751e-11 -6.802422
m_suppression_organes_touche=s_m_suppression_organes_touche_0 0.00000 0.000000 36.95652 1.028751e-11 -6.802422
> hcpc$desc.var$category$`2`[c(1:10,(nrow(hcpc$desc.var$category$`2`)-10):(nrow(hcpc$desc.var$category$`2`))),]
Cla/ModMod/Cla Global p.valuev.test
syst_req=REF 100.000000 94.444444 36.95652 1.028751e-11 6.802422
r_produits_biocontrôle=r_produits_biocontrôle_0 100.000000 94.444444 36.95652 1.028751e-11 6.802422
m_suppression_organes_touche=s_m_suppression_organes_touche_0 100.000000 94.444444 36.95652 1.028751e-11 6.802422
r_vigueur=r_vigueur_0 100.000000 83.333333 32.60870 1.594564e-09 6.034499
m_reduction_dose=m_reduction_dose_0 89.473684 94.444444 41.30435 1.651322e-09 6.028848
a_herbicide=a_herbicide_1 89.473684 94.444444 41.30435 1.651322e-09 6.028848
r_prophylaxie=r_prophylaxie_0 85.000000 94.444444 43.47826 1.064934e-08 5.720049
```

```

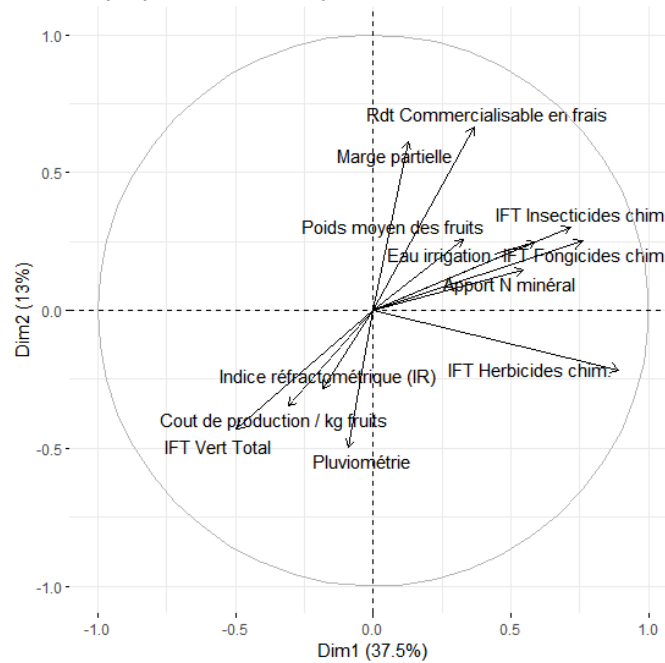
m_vigueur=m_vigueur_0      85.000000 94.444444 43.47826 1.064934e-08 5.720049
m_gestion_irrigation=m_gestion_irrigation_0 73.913043 94.444444 50.00000 8.475058e-07 4.924097
m_aeration=m_aeration_0    100.000000 61.111111 23.91304 2.385467e-06 4.717682
a_travail_sol=a_travail_sol_1 0.000000 0.000000 28.26087 3.679232e-04 -3.562084
syst_req=ECO                4.545455 5.555556 47.82609 2.796598e-06 -4.685220
m_aeration=m_aeration_1    20.000000 38.888889 76.08696 2.385467e-06 -4.717682
m_gestion_irrigation=m_gestion_irrigation_1 4.347826 5.555556 50.00000 8.475058e-07 -4.924097
r_prophylaxie=r_prophylaxie_1 3.846154 5.555556 56.52174 1.064934e-08 -5.720049
m_vigueur=m_vigueur_1      3.846154 5.555556 56.52174 1.064934e-08 -5.720049
m_reduction_dose=m_reduction_dose_1 3.703704 5.555556 58.69565 1.651322e-09 -6.028848
a_herbicide=a_herbicide_0   3.703704 5.555556 58.69565 1.651322e-09 -6.028848
r_vigueur=r_vigueur_1       9.677419 16.666667 67.39130 1.594564e-09 -6.034499
r_produits_biocontrôle=r_produits_biocontrôle_1 3.448276 5.555556 63.04348 1.028751e-11 -6.802422
m_suppression_organes_touchees=m_suppression_organes_touchees_1 3.448276 5.555556 63.04348 1.028751e-11 -6.802422
> #hcpc$desc.var$category$`3`[c(1:5,(nrow(hcpc$desc.var$category$`3`)-5):(nrow(hcpc$desc.var$category$`3`))),]
>
> # Description par les individus
>hcpc$desc.ind$para
Cluster: 1
      11-17      12-17      6-16      6-17      6-18
0.2560241 0.2560241 0.4859513 0.4859513 0.4859513
-----
Cluster: 2
      13-16      13-17      16-16      16-17      16-18
0.5019713 0.5019713 0.5019713 0.5019713 0.5019713

```

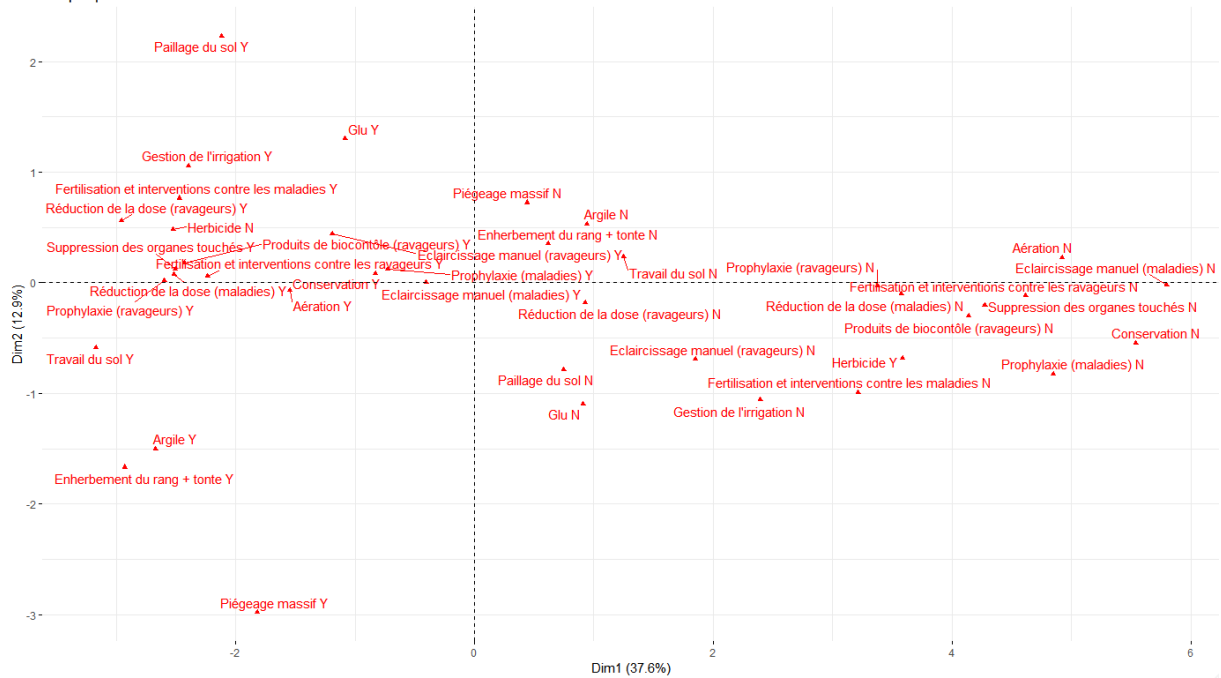
Annexe 21 : AFDM des leviers et variables agronomiques – Graphiques de l'ensemble des variables, des variables quantitatives, des modalités, des individus, sortie de l'HCPC et caractéristiques des clusters

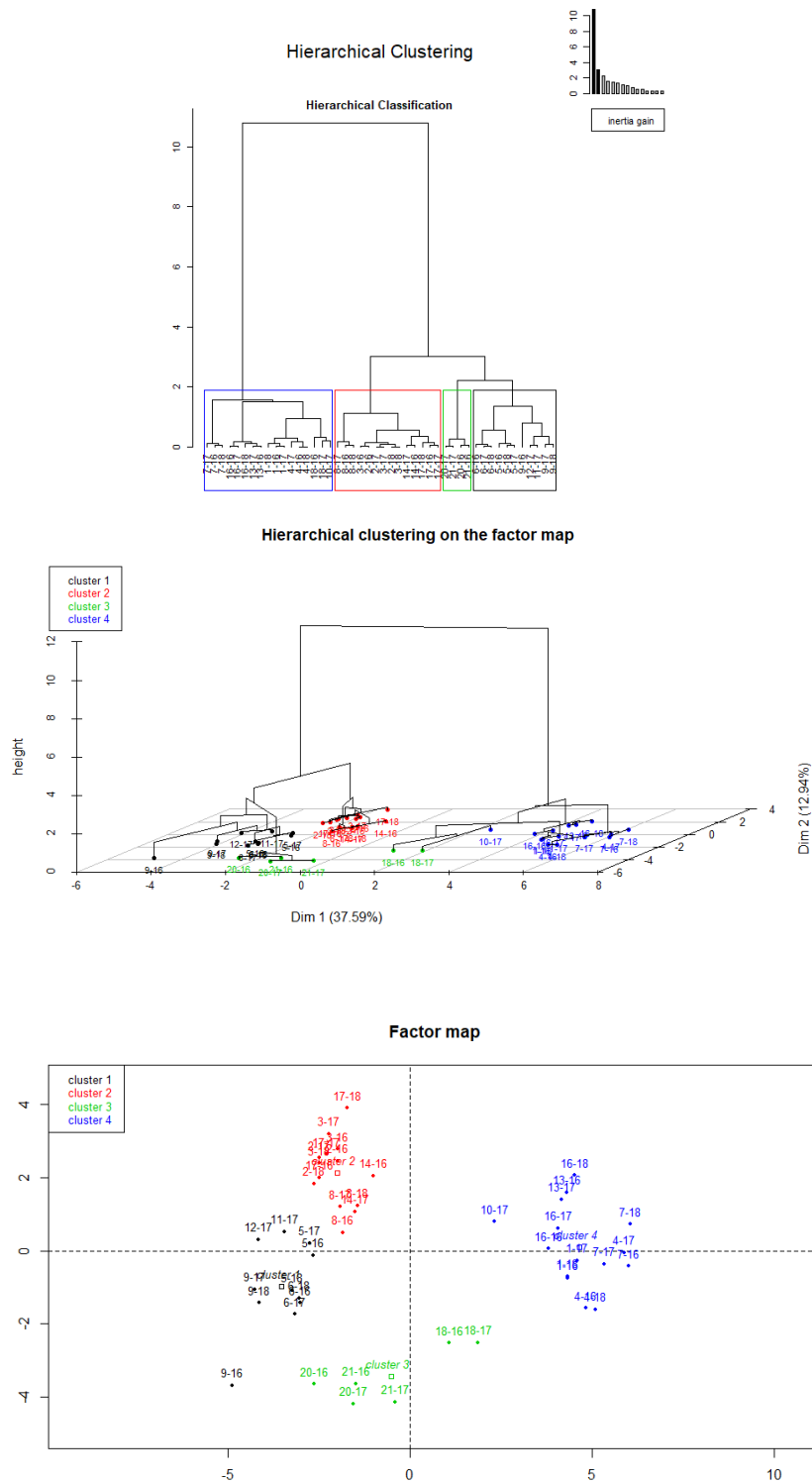


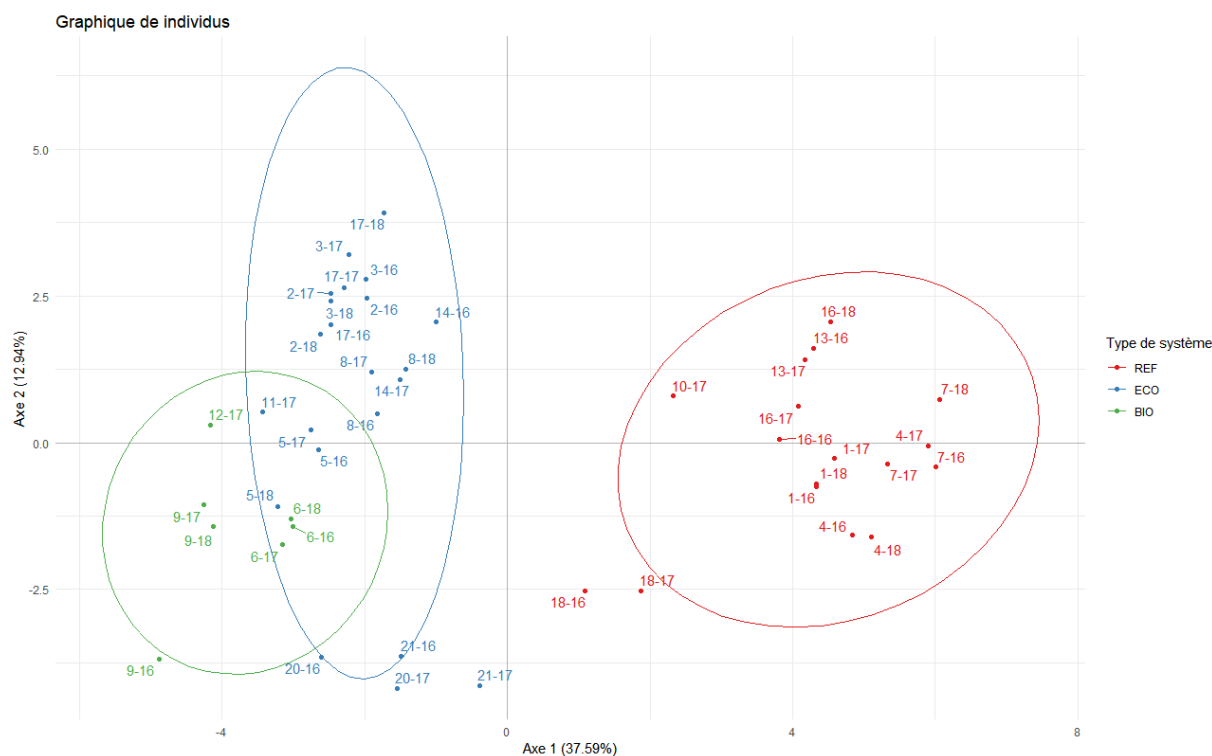
Graphique des variables quantitatives



Graphique des modalités







```
>hpcpc$desc.var$quanti
```

```
$`1`
```

```
v.testMean in categoryOverallmeansd in categoryOverallsdp.value
```

| | | | | | | |
|--------------------------------|-----------|--------------|--------------|-------------|--------------|--------------|
| Cout.de.production./.kg.fruits | 2.454078 | 0.8486230 | 0.5482663 | 0.8547248 | 0.4602750 | 1.412463e-02 |
| IFT.Vert.Total | 2.136968 | 7.9390909 | 5.8708696 | 4.1457655 | 3.6397163 | 3.260058e-02 |
| Rdt.Commercialisable.en.frais | -2.363852 | 18.3363636 | 27.1735057 | 12.4422734 | 14.0591870 | 1.808606e-02 |
| Eau.irrigation | -2.653031 | 4276.9090909 | 5448.4347826 | 851.2670187 | 1660.6501562 | 7.977266e-03 |
| IFT.Herbicides.chim. | -2.910033 | 0.0000000 | 0.5421467 | 0.0000000 | 0.7006281 | 3.613907e-03 |
| IFT.Fongicides.chim. | -3.011971 | 5.2590909 | 8.2429028 | 2.5545697 | 3.7255401 | 2.595575e-03 |
| Poids.moyen.des.fruits | -3.260495 | 154.1000000 | 173.2898535 | 22.4538719 | 22.1338388 | 1.112181e-03 |
| IFT.Insecticides.chim. | -3.729132 | 0.2727273 | 4.6699275 | 0.6165755 | 4.4344228 | 1.921407e-04 |
| Apport.N.minéral | -5.427036 | 45.5454545 | 114.6739130 | 44.1451400 | 47.9030063 | 5.729767e-08 |

```
$`2`
```

```
v.testMean in categoryOverallmeansd in categoryOverallsdp.value
```

| | | | | | | |
|----------------------|-----------|---|-----------|---|-----------|--------------|
| IFT.Herbicides.chim. | -3.433402 | 0 | 0.5421467 | 0 | 0.7006281 | 0.0005960583 |
|----------------------|-----------|---|-----------|---|-----------|--------------|

```
$`3`
```

```
v.testMean in categoryOverallmeansd in categoryOverallsdp.value
```

| | | | | | | |
|-------------------------------|-----------|-------------|-------------|--------------|--------------|--------------|
| Pluviométrie | 3.658664 | 4362.500000 | 2520.06522 | 2137.5000000 | 1308.342486 | 0.0002535332 |
| IFT.Vert.Total | 2.780870 | 9.766667 | 5.87087 | 3.8525605 | 3.639716 | 0.0054213442 |
| Indice.réfractométrique.(IR) | 2.631482 | 13.800000 | 12.17500 | 0.6198118 | 1.604374 | 0.0085013468 |
| Apport.N.minéral | 2.187133 | 155.000000 | 114.67391 | 15.0000000 | 47.903006 | 0.0287328143 |
| Eau.irrigation | -2.266067 | 4000.000000 | 5448.43478 | 0.0000000 | 1660.650156 | 0.0234473022 |
| Marge partielle | -2.534114 | 7963.461428 | 20325.08711 | 4449.5933168 | 12673.640742 | 0.0112732173 |
| Rdt.Commercialisable.en.frais | -2.568691 | 13.273333 | 27.17351 | 5.4110096 | 14.059187 | 0.0102083434 |

```
$`4`
```

```
v.testMean in categoryOverallmeansd in categoryOverallsdp.value
```

| | | | | | | |
|-------------------------------|----------|-------------|--------------|-------------|--------------|--------------|
| IFT.Herbicides.chim. | 5.779471 | 1.409917 | 0.5421467 | 0.423340 | 0.7006281 | 7.493572e-09 |
| IFT.Insecticides.chim. | 4.598334 | 9.039778 | 4.6699275 | 4.048063 | 4.4344228 | 4.258832e-06 |
| IFT.Fongicides.chim. | 4.551987 | 11.877190 | 8.2429028 | 2.269660 | 3.7255401 | 5.314170e-06 |
| Eau.irrigation | 3.897626 | 6835.533333 | 5448.4347826 | 1630.647473 | 1660.6501562 | 9.714013e-05 |
| Apport.N.minéral | 2.713816 | 142.533333 | 114.6739130 | 18.604181 | 47.9030063 | 6.651317e-03 |
| Rdt.Commercialisable.en.frais | 2.555979 | 34.874495 | 27.1735057 | 12.206000 | 14.0591870 | 1.058894e-02 |

| | | | | | | |
|----------------|-----------|----------|-----------|----------|-----------|--------------|
| IFT.Vert.Total | -3.112212 | 3.443333 | 5.8708696 | 1.475695 | 3.6397163 | 1.856908e-03 |
|----------------|-----------|----------|-----------|----------|-----------|--------------|

```
> hcpc$desc.var$category$`1`[,c(1,2,4,5)]
```

```
Clap/ModMod/Clap.valuev.test
```

| | | | | |
|---|-----------|-----------|--------------|-----------|
| Travail.du.sol=Travail du sol Y | 84.61538 | 100.00000 | 5.846733e-09 | 5.821084 |
| Type.de.système=BIO | 100.00000 | 63.63636 | 6.165380e-06 | 4.520637 |
| Prophylaxie.(ravageurs)=Prophylaxie (ravageurs) Y | 42.30769 | 100.00000 | 5.791384e-04 | 3.441201 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies Y | 42.30769 | 100.00000 | 5.791384e-04 | 3.441201 |
| Réduction.de.la.dose.(maladies)=Réduction de la dose (maladies) Y | 40.74074 | 100.00000 | 9.772961e-04 | 3.296982 |
| Herbicide=Herbicide N | 40.74074 | 100.00000 | 9.772961e-04 | 3.296982 |
| Enherbement.du.rang.+.tonte=Enherbement du rang + tonte Y | 75.00000 | 54.54545 | 1.143283e-03 | 3.252665 |
| Produits.de.biocontôle.(ravageurs)=Produits de biocontôle (ravageurs) Y | 37.93103 | 100.00000 | 2.593348e-03 | 3.012231 |
| Suppression.des.organes.touchés=Suppression des organes touchés Y | 37.93103 | 100.00000 | 2.593348e-03 | 3.012231 |
| Fertilisation.et.interventions.contre.les.ravageurs=Fertilisation et interventions contre les ravageurs Y | 35.48387 | 100.00000 | 6.346877e-03 | 2.729301 |
| Paillage.du.sol=Paillage du sol N | 32.35294 | 100.00000 | 2.144535e-02 | 2.300051 |
| Aération=Aération Y | 31.42857 | 100.00000 | 3.127447e-02 | 2.153563 |
| Aération=Aération N | 0.00000 | 0.00000 | 3.127447e-02 | -2.153563 |
| Paillage.du.sol=Paillage du sol Y | 0.00000 | 0.00000 | 2.144535e-02 | -2.300051 |
| Fertilisation.et.interventions.contre.les.ravageurs=Fertilisation et interventions contre les ravageurs N | 0.00000 | 0.00000 | 6.346877e-03 | -2.729301 |
| Type.de.système=REF | 0.00000 | 0.00000 | 2.593348e-03 | -3.012231 |
| Produits.de.biocontôle.(ravageurs)=Produits de biocontôle (ravageurs) N | 0.00000 | 0.00000 | 2.593348e-03 | -3.012231 |
| Suppression.des.organes.touchés=Suppression des organes touchés N | 0.00000 | 0.00000 | 2.593348e-03 | -3.012231 |
| Enherbement.du.rang.+.tonte=Enherbement du rang + tonte N | 13.15789 | 45.45455 | 1.143283e-03 | -3.252665 |
| Réduction.de.la.dose.(maladies)=Réduction de la dose (maladies) N | 0.00000 | 0.00000 | 9.772961e-04 | -3.296982 |
| Herbicide=Herbicide Y | 0.00000 | 0.00000 | 9.772961e-04 | -3.296982 |
| Prophylaxie.(ravageurs)=Prophylaxie (ravageurs) N | 0.00000 | 0.00000 | 5.791384e-04 | -3.441201 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies N | 0.00000 | 0.00000 | 5.791384e-04 | -3.441201 |
| Travail.du.sol=Travail du sol N | 0.00000 | 0.00000 | 5.846733e-09 | -5.821084 |

```
> hcpc$desc.var$category$`2`[,c(1,2,4,5)]
```

```
Clap/ModMod/Clap.valuev.test
```

| | | | | |
|---|------------|-----------|--------------|-----------|
| Paillage.du.sol=Paillage du sol Y | 100.000000 | 85.71429 | 2.338693e-09 | 5.972344 |
| Type.de.système=ECO | 63.636364 | 100.00000 | 1.333055e-06 | 4.834761 |
| Gestion.de.l'irrigation=Gestion de l'irrigation Y | 60.869565 | 100.00000 | 3.406697e-06 | 4.644640 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies Y | 53.846154 | 100.00000 | 4.026096e-05 | 4.105977 |
| Herbicide=Herbicide N | 51.851852 | 100.00000 | 8.361892e-05 | 3.933784 |
| Produits.de.biocontôle.(ravageurs)=Produits de biocontôle (ravageurs) Y | 48.275862 | 100.00000 | 3.233265e-04 | 3.595856 |
| Suppression.des.organes.touchés=Suppression des organes touchés Y | 48.275862 | 100.00000 | 3.233265e-04 | 3.595856 |
| Fertilisation.et.interventions.contre.les.ravageurs=Fertilisation et interventions contre les ravageurs Y | 45.161290 | 100.00000 | 1.105491e-03 | 3.262205 |
| Glu=Glu Y | 52.380952 | 78.57143 | 4.160506e-03 | 2.865729 |
| Aération=Aération Y | 40.000000 | 100.00000 | 9.671432e-03 | 2.587361 |
| Réduction.de.la.dose.(maladies)=Réduction de la dose (maladies) Y | 44.444444 | 85.71429 | 1.573708e-02 | 2.414957 |
| Piégeage.massif=Piégeage massif N | 37.837838 | 100.00000 | 2.545919e-02 | 2.234363 |
| Enherbement.du.rang.+.tonte=Enherbement du rang + tonte N | 36.842105 | 100.00000 | 4.031038e-02 | 2.050554 |
| Enherbement.du.rang.+.tonte=Enherbement du rang + tonte Y | 0.000000 | 0.00000 | 4.031038e-02 | -2.050554 |
| Piégeage.massif=Piégeage massif Y | 0.000000 | 0.00000 | 2.545919e-02 | -2.234363 |
| Réduction.de.la.dose.(maladies)=Réduction de la dose (maladies) N | 10.526316 | 14.28571 | 1.573708e-02 | -2.414957 |
| Aération=Aération N | 0.000000 | 0.00000 | 9.671432e-03 | -2.587361 |
| Glu=Glu N | 12.000000 | 21.42857 | 4.160506e-03 | -2.865729 |
| Fertilisation.et.interventions.contre.les.ravageurs=Fertilisation et interventions contre les ravageurs N | 0.000000 | 0.00000 | 1.105491e-03 | -3.262205 |
| Type.de.système=REF | 0.000000 | 0.00000 | 3.233265e-04 | -3.595856 |
| Produits.de.biocontôle.(ravageurs)=Produits de biocontôle (ravageurs) N | 0.000000 | 0.00000 | 3.233265e-04 | -3.595856 |
| Suppression.des.organes.touchés=Suppression des organes touchés N | 0.000000 | 0.00000 | 3.233265e-04 | -3.595856 |
| Herbicide=Herbicide Y | 0.000000 | 0.00000 | 8.361892e-05 | -3.933784 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies N | 0.000000 | 0.00000 | 4.026096e-05 | -4.105977 |
| Gestion.de.l'irrigation=Gestion de l'irrigation N | 0.000000 | 0.00000 | 3.406697e-06 | -4.644640 |
| Paillage.du.sol=Paillage du sol N | 5.882353 | 14.28571 | 2.338693e-09 | -5.972344 |

```
> hcpc$desc.var$category$`3`[,c(1,2,4,5)]
```

```
Clap/ModMod/Clap.valuev.test
```

| | | | | |
|---|-----------|-----------|--------------|----------|
| Piégeage.massif=Piégeage massif Y | 66.666667 | 100.00000 | 8.967826e-06 | 4.440672 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies N | 30.000000 | 100.00000 | 4.138011e-03 | 2.867445 |
| Gestion.de.l'irrigation=Gestion de l'irrigation N | 26.086957 | 100.00000 | 1.077708e-02 | 2.549845 |

| | | | | |
|---|------------|------------|--------------|-----------|
| Glu=Glu N | 24.000000 | 100.00000 | 1.890717e-02 | 2.347356 |
| Argile=Argile Y | 33.333333 | 66.66667 | 3.559362e-02 | 2.101540 |
| Argile=Argile N | 5.882353 | 33.33333 | 3.559362e-02 | -2.101540 |
| Glu=Glu Y | 0.000000 | 0.00000 | 1.890717e-02 | -2.347356 |
| Gestion.de.l'irrigation=Gestion de l'irrigation Y | 0.000000 | 0.00000 | 1.077708e-02 | -2.549845 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies Y | 0.000000 | 0.00000 | 4.138011e-03 | -2.867445 |
| Piégeage.massif=Piégeage massif N | 0.000000 | 0.00000 | 8.967826e-06 | -4.440672 |
| > hcpc\$desc.var\$category\$`4`[,c(1,2,4,5)] | | | | |
| Clap/ModMod/Clap.valuev.test | | | | |
| Type.de.système=REF | 88.235294 | 100.000000 | 2.657606e-10 | 6.317539 |
| Produits.de.biocontrôle.(ravageurs)=Produits de biocontrôle (ravageurs) N | 88.235294 | 100.000000 | 2.657606e-10 | 6.317539 |
| Suppression.des.organes.touchés=Suppression des organes touchés N | 88.235294 | 100.000000 | 2.657606e-10 | 6.317539 |
| Fertilisation.et.interventions.contre.les.ravageurs=Fertilisation et interventions contre les ravageurs N | 93.333333 | 93.333333 | 9.125750e-10 | 6.123995 |
| Réduction.de.la.dose.(maladies)=Réduction de la dose (maladies) N | 78.947368 | 100.000000 | 7.574177e-09 | 5.777671 |
| Herbicide=Herbicide Y | 78.947368 | 100.000000 | 7.574177e-09 | 5.777671 |
| Prophylaxie.(ravageurs)=Prophylaxie (ravageurs) N | 75.000000 | 100.000000 | 3.029671e-08 | 5.539713 |
| Aération=Aération N | 100.000000 | 73.333333 | 1.023178e-07 | 5.322559 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies N | 70.000000 | 93.333333 | 2.029879e-06 | 4.750427 |
| Gestion.de.l'irrigation=Gestion de l'irrigation N | 60.869565 | 93.333333 | 3.864471e-05 | 4.115437 |
| Conservation=Conservation N | 100.000000 | 40.000000 | 5.343329e-04 | 3.462927 |
| Prophylaxie.(maladies)=Prophylaxie (maladies) N | 100.000000 | 40.000000 | 5.343329e-04 | 3.462927 |
| Travail.du.sol=Travail du sol N | 45.454545 | 100.000000 | 2.026734e-03 | 3.086287 |
| Argile=Argile N | 44.117647 | 100.000000 | 3.626787e-03 | 2.908921 |
| Paillage.du.sol=Paillage du sol N | 44.117647 | 100.000000 | 3.626787e-03 | 2.908921 |
| Réduction.de.la.dose.(ravageurs)=Réduction de la dose (ravageurs) N | 42.857143 | 100.000000 | 6.346877e-03 | 2.729301 |
| Eclaircissage.manuel.(ravageurs)=Eclaircissage manuel (ravageurs) N | 55.555556 | 66.666667 | 1.120105e-02 | 2.536363 |
| Piégeage.massif=Piégeage massif N | 40.540541 | 100.000000 | 1.829879e-02 | 2.359516 |
| Eclaircissage.manuel.(maladies)=Eclaircissage manuel (maladies) N | 100.000000 | 20.000000 | 2.997365e-02 | 2.170438 |
| Enherbement.du.rang.+.tonte=Enherbement du rang + tonte N | 39.473684 | 100.000000 | 3.023278e-02 | 2.167027 |
| Type.de.système=BIO | 0.000000 | 0.000000 | 4.912827e-02 | -1.967477 |
| Enherbement.du.rang.+.tonte=Enherbement du rang + tonte Y | 0.000000 | 0.000000 | 3.023278e-02 | -2.167027 |
| Eclaircissage.manuel.(maladies)=Eclaircissage manuel (maladies) Y | 27.906977 | 80.000000 | 2.997365e-02 | -2.170438 |
| Piégeage.massif=Piégeage massif Y | 0.000000 | 0.000000 | 1.829879e-02 | -2.359516 |
| Eclaircissage.manuel.(ravageurs)=Eclaircissage manuel (ravageurs) Y | 17.857143 | 33.333333 | 1.120105e-02 | -2.536363 |
| Réduction.de.la.dose.(ravageurs)=Réduction de la dose (ravageurs) Y | 0.000000 | 0.000000 | 6.346877e-03 | -2.729301 |
| Argile=Argile Y | 0.000000 | 0.000000 | 3.626787e-03 | -2.908921 |
| Paillage.du.sol=Paillage du sol Y | 0.000000 | 0.000000 | 3.626787e-03 | -2.908921 |
| Travail.du.sol=Travail du sol Y | 0.000000 | 0.000000 | 2.026734e-03 | -3.086287 |
| Conservation=Conservation Y | 22.500000 | 60.000000 | 5.343329e-04 | -3.462927 |
| Prophylaxie.(maladies)=Prophylaxie (maladies) Y | 22.500000 | 60.000000 | 5.343329e-04 | -3.462927 |
| Gestion.de.l'irrigation=Gestion de l'irrigation Y | 4.347826 | 6.666667 | 3.864471e-05 | -4.115437 |
| Type.de.système=ECO | 0.000000 | 0.000000 | 2.555022e-06 | -4.703689 |
| Fertilisation.et.interventions.contre.les.maladies=Fertilisation et interventions contre les maladies Y | 3.846154 | 6.666667 | 2.029879e-06 | -4.750427 |
| Aération=Aération Y | 11.428571 | 26.666667 | 1.023178e-07 | -5.322559 |
| Prophylaxie.(ravageurs)=Prophylaxie (ravageurs) Y | 0.000000 | 0.000000 | 3.029671e-08 | -5.539713 |
| Réduction.de.la.dose.(maladies)=Réduction de la dose (maladies) Y | 0.000000 | 0.000000 | 7.574177e-09 | -5.777671 |
| Herbicide=Herbicide N | 0.000000 | 0.000000 | 7.574177e-09 | -5.777671 |
| Fertilisation.et.interventions.contre.les.ravageurs=Fertilisation et interventions contre les ravageurs Y | 3.225806 | 6.666667 | 9.125750e-10 | -6.123995 |
| Produits.de.biocontrôle.(ravageurs)=Produits de biocontrôle (ravageurs) Y | 0.000000 | 0.000000 | 2.657606e-10 | -6.317539 |
| Suppression.des.organes.touchés=Suppression des organes touchés Y | 0.000000 | 0.000000 | 2.657606e-10 | -6.317539 |

>hcpc\$desc.ind

\$para

Cluster: 1

| | | | | |
|----------|----------|----------|----------|----------|
| 11-17 | 5-18 | 5-17 | 6-16 | 6-17 |
| 2.597891 | 2.922253 | 3.148373 | 3.229606 | 3.240358 |

Cluster: 2

| | | | | |
|----------|----------|----------|----------|----------|
| 17-16 | 3-16 | 17-17 | 2-16 | 2-18 |
| 1.698240 | 1.874386 | 2.173640 | 2.345910 | 2.433802 |

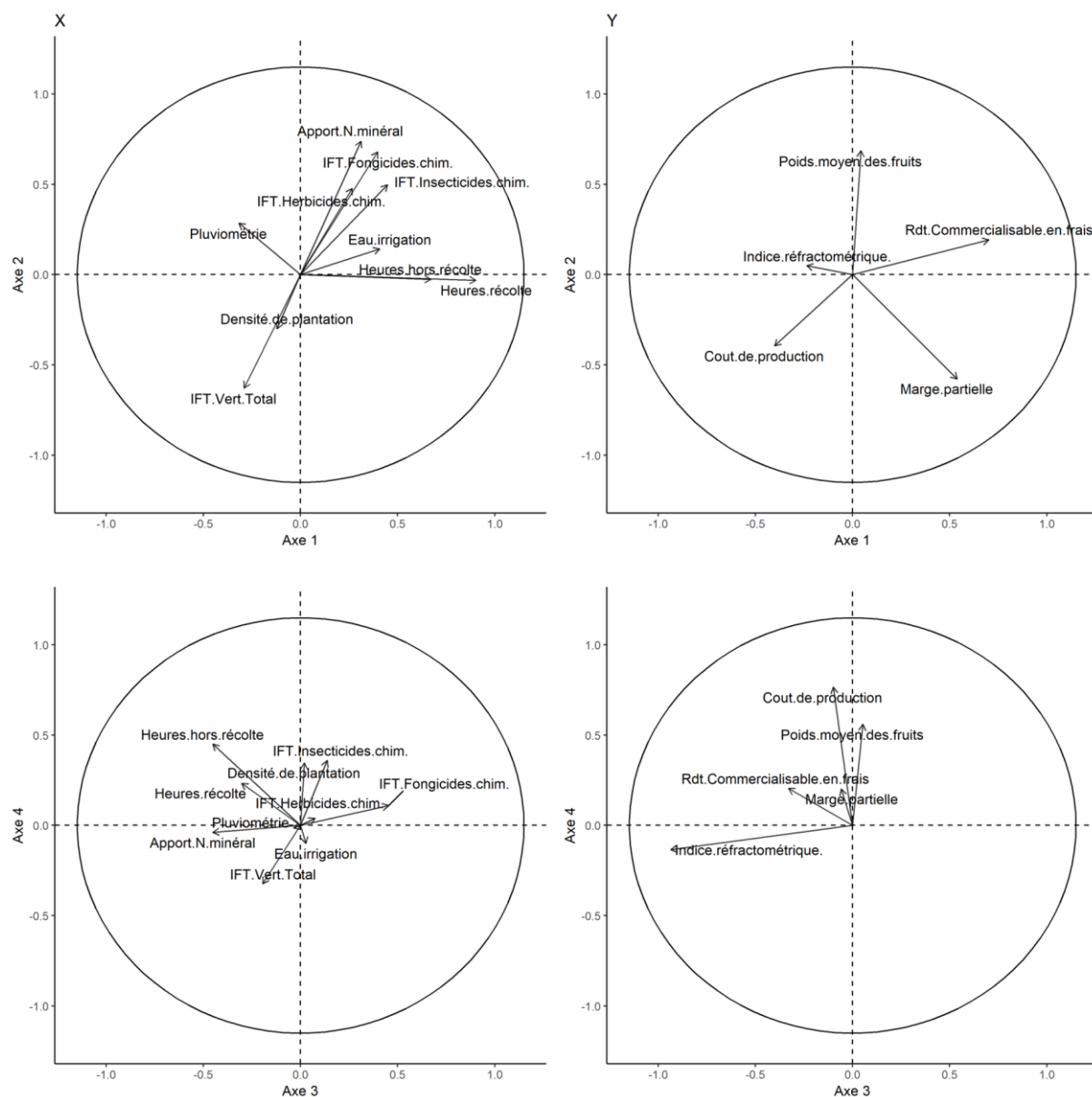
Cluster: 3

| 21-16 | 21-17 | 20-17 | 20-16 | 18-17 |
|----------|----------|----------|----------|----------|
| 2.183043 | 2.315092 | 2.839909 | 3.010750 | 3.747312 |

Cluster: 4

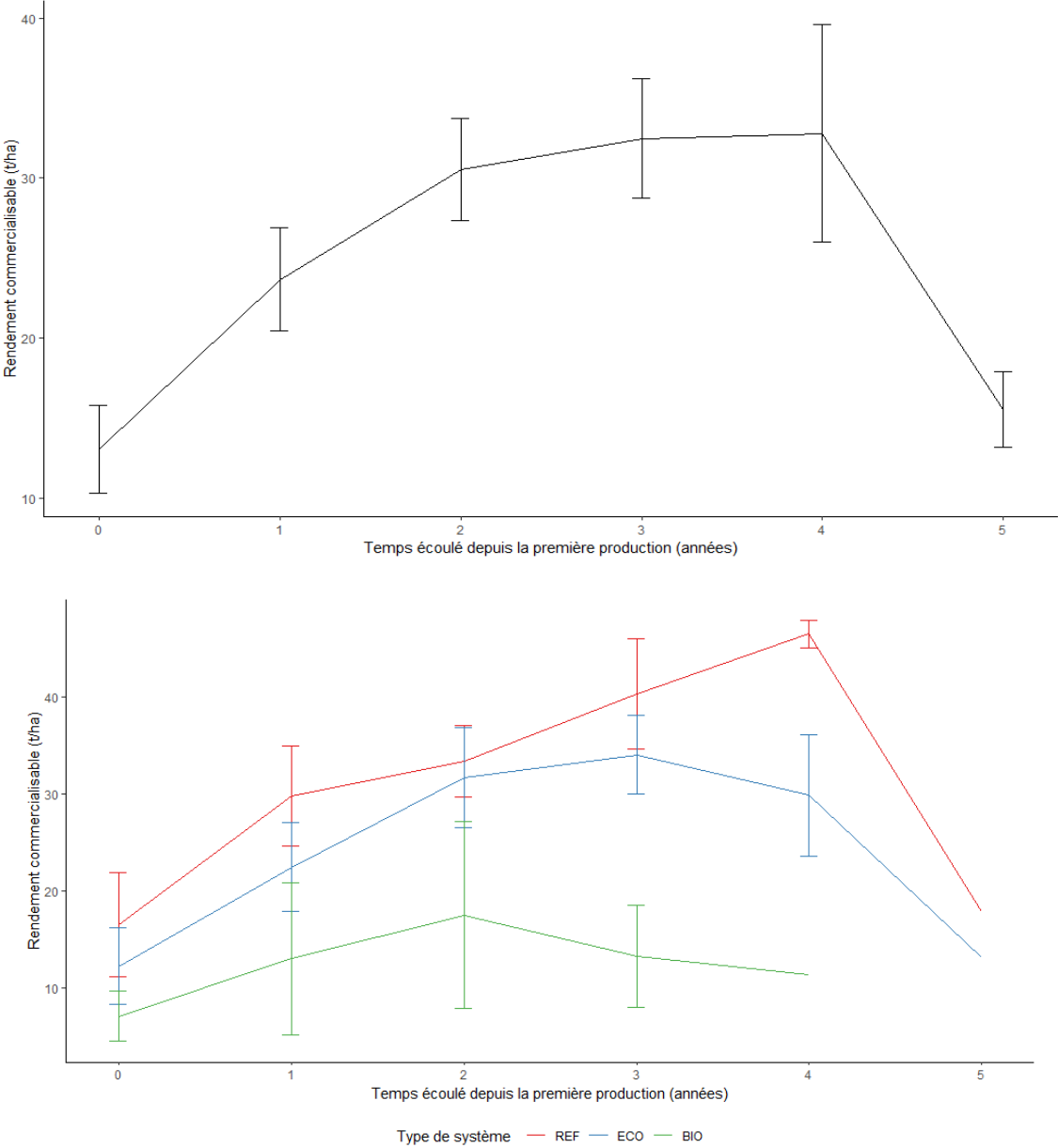
| 4-16 | 4-18 | 1-16 | 1-17 | 13-17 |
|----------|----------|----------|----------|----------|
| 2.981850 | 3.128712 | 3.163290 | 3.303071 | 3.305993 |

Annexe 22 : ACPVI des performances et des variables explicatives

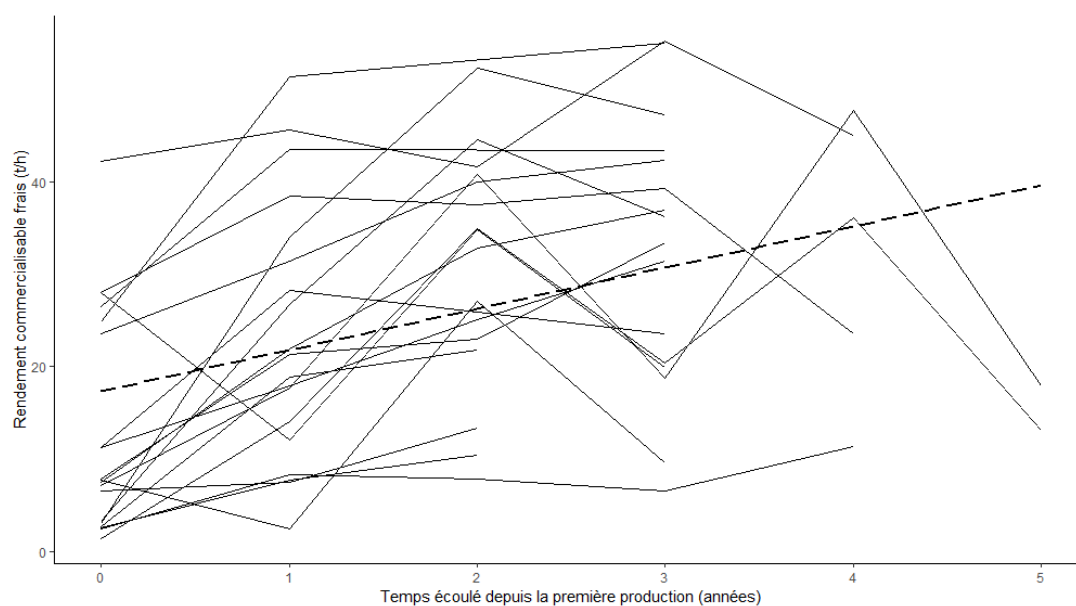


La prise en compte des 3 premiers axes permet d'atteindre 95.02% de l'inertie totale.

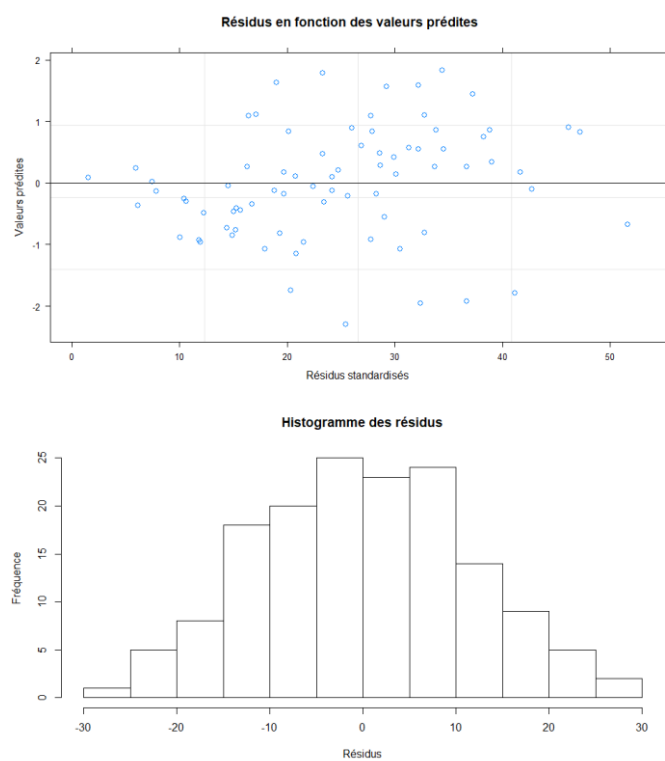
Annexe 23 : Evolution des rendements commercialisables (euros/ha) globale et par type de système

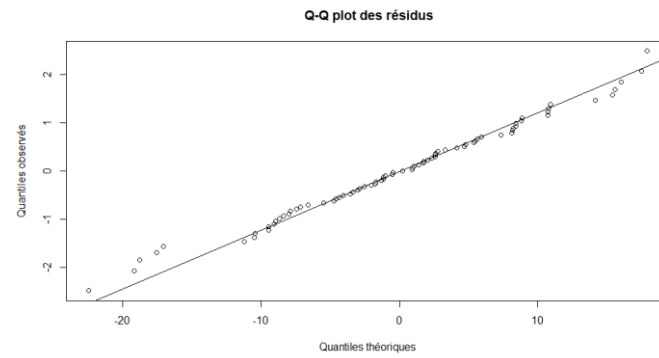


Annexe 24 : Evolution moyenne des rendements commercialisables (euros/ha) prédits par le modèle 1

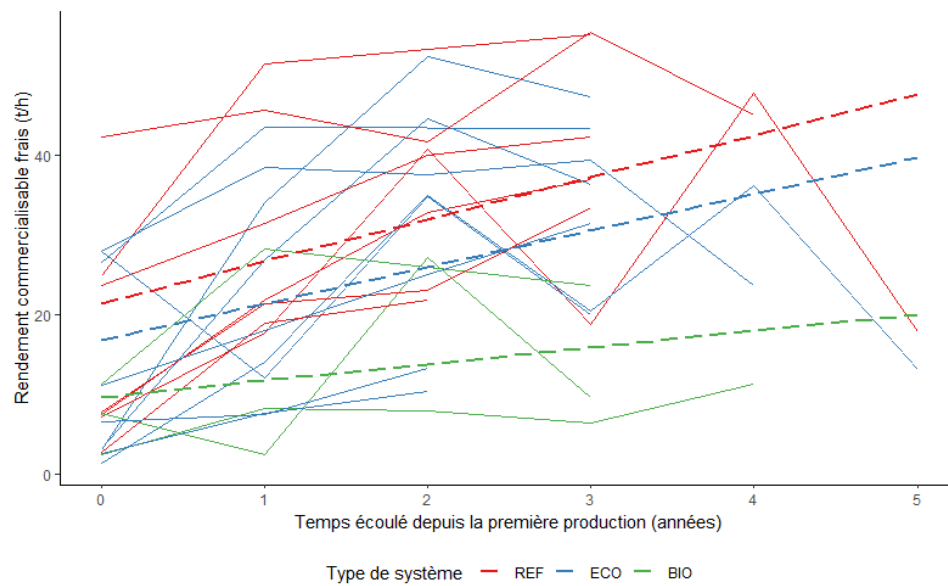


Annexe 25 : Analyse des résidus - Rendements commercialisables - Modèle 1

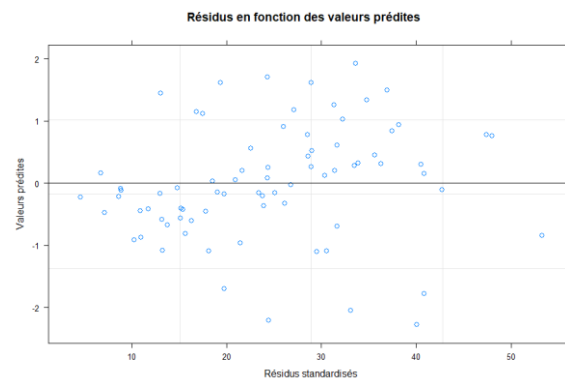


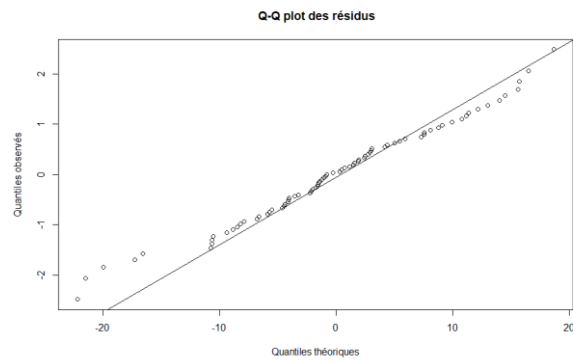
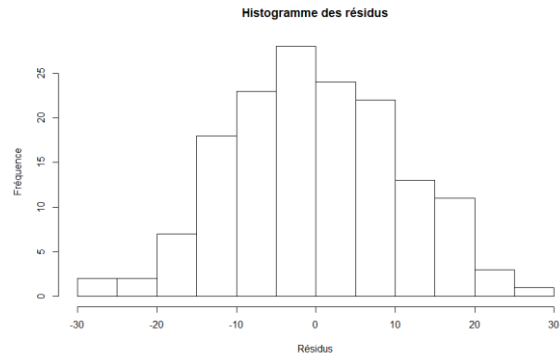


Annexe 26 : Evolution moyenne des rendements commercialisables (euros/ha) par type des systèmes prédite par le modèle 2

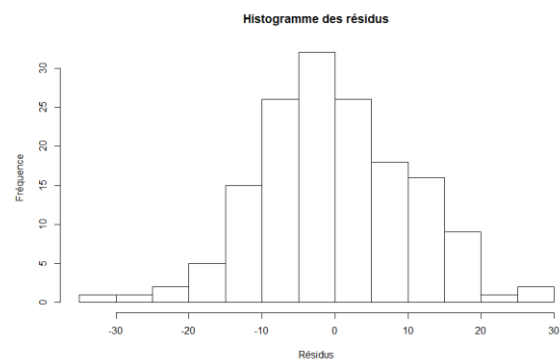
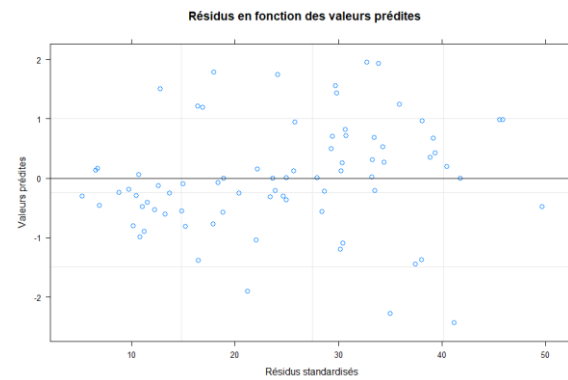


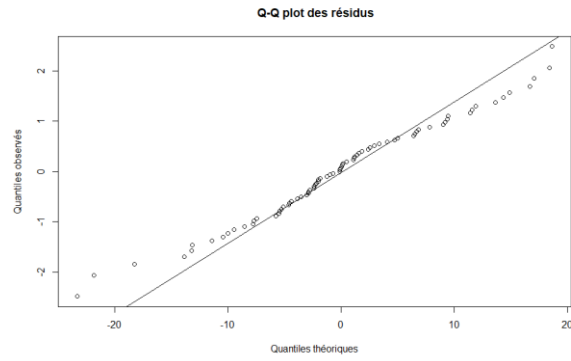
Annexe 27 : Analyse des résidus - Rendements commercialisables - Modèle 2





Annexe 28 : Analyse des résidus - Rendements commercialisables - Modèle 3





Annexe 29 : Estimation des effets fixes, des effets aléatoires et résultats du test du rapport de vraisemblance sur les modèles - Modèle mixte sur le rendement commercialisable

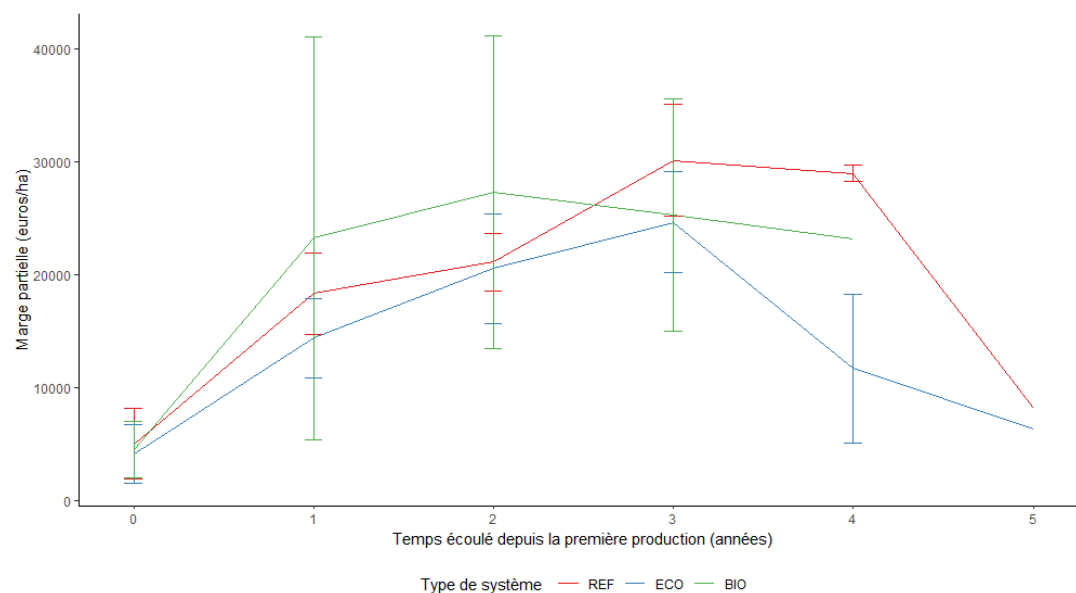
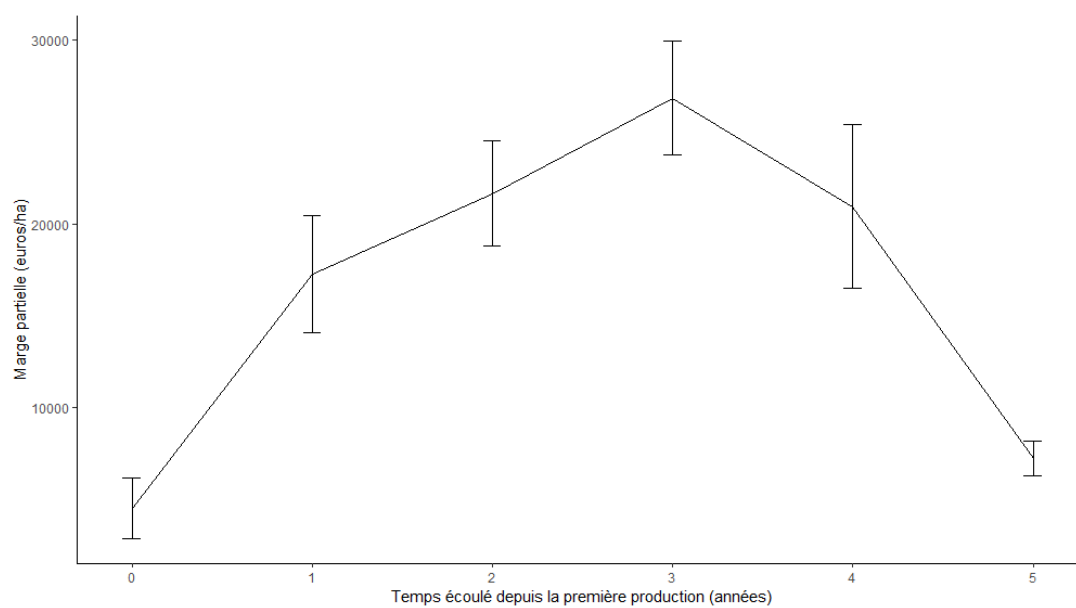
| M1 | | | | | M2 | | | M3 | | |
|---------------------|-----------|------------|-------|--------|------------|-------|--------|------------|-------|-------|
| Variables | Param. | Estimation | Std | P* | Estimation | Std | P* | Estimation | Std | P* |
| Intercept | β_0 | 17.362 | 2.926 | <0.001 | 21.408 | 4.35 | <0.001 | 4.129 | 9.359 | 0.661 |
| Temps | β_1 | 4.462 | 0.878 | <0.001 | 5.251 | 1.42 | 0.001 | 10.419 | 4.593 | 0.027 |
| Système (ECO) | β_2 | | | | -4.566 | 5.812 | 0.444 | 0.988 | 6.205 | 0.875 |
| Système (BIO) | β_3 | | | | -11.767 | 8.009 | 0.161 | -5.115 | 8.354 | 0.549 |
| Temps*système (ECO) | β_4 | | | | -0.677 | 1.943 | 0.729 | -2.394 | 2.580 | 0.358 |
| Temps*système (BIO) | β_5 | | | | -3.165 | 2.674 | 0.242 | -5.017 | 3.160 | 0.118 |
| IFT | β_6 | | | | | | | 0.715 | 0.346 | 0.044 |
| Temps*IFT | β_7 | | | | | | | -0.215 | 0.176 | 0.228 |

* Test de Wald

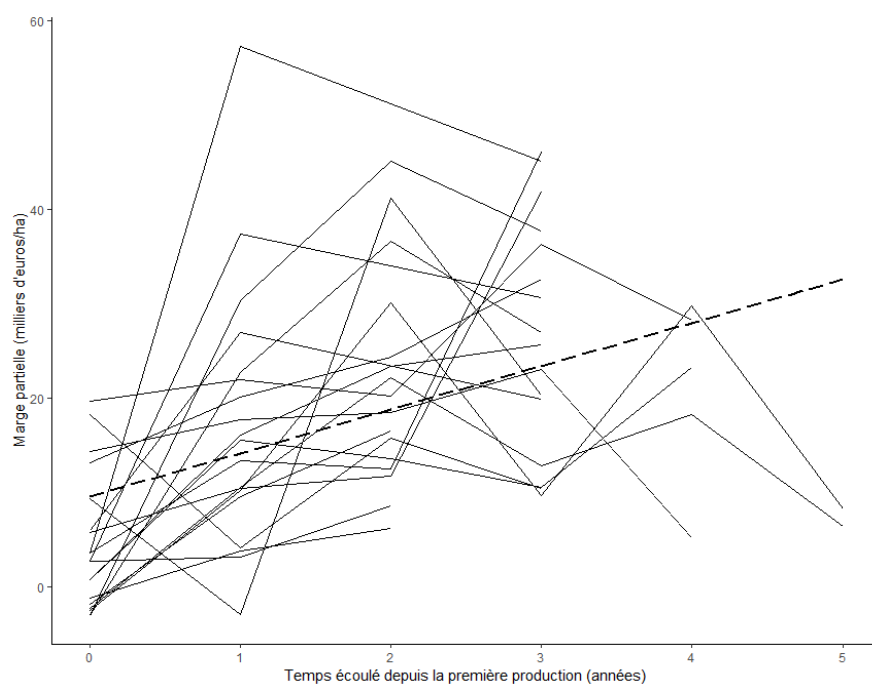
| M1 | | | M2 | | M3 | |
|---------------------|------------|-------|------------|-------|------------|-------|
| Paramètre | Estimation | Std | Estimation | Std | Estimation | Std |
| σ_b^2 | 97.439 | 9.871 | 63.177 | 7.948 | 52.606 | 7.253 |
| σ_ϵ^2 | 95.88 | 9.792 | 94.381 | 9.715 | 91.346 | 9.558 |

| Variables | M1 | M2 | M3 |
|---------------|--------|--------|--------|
| Intercept | <0.001 | <0.001 | <0.001 |
| Temps | <0.001 | <0.001 | <0.001 |
| Système | | 0.068 | 0,227 |
| Temps*système | | 0.497 | 0,292 |
| IFT | | | 0.012 |
| Temps*IFT | | | 0,603 |

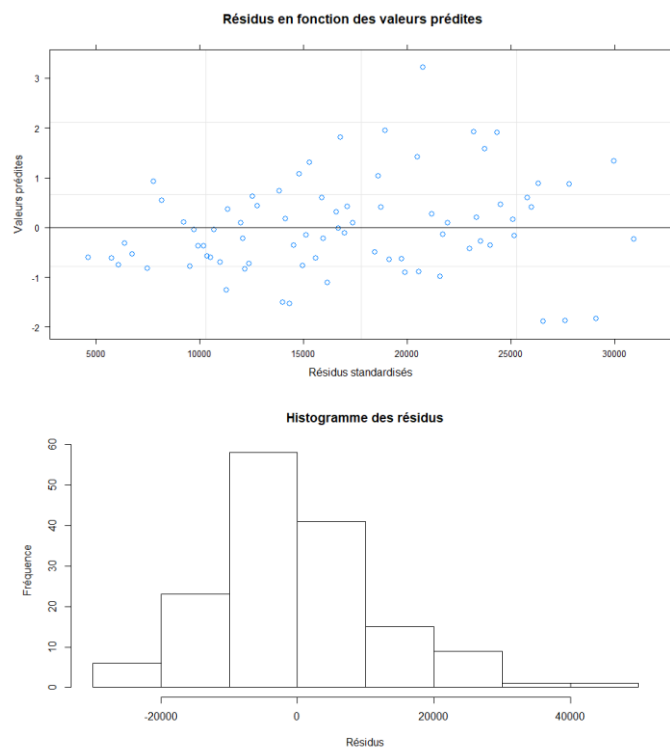
Annexe 30 : Evolution de la marge partielle (euros/ha) globale et par type de système

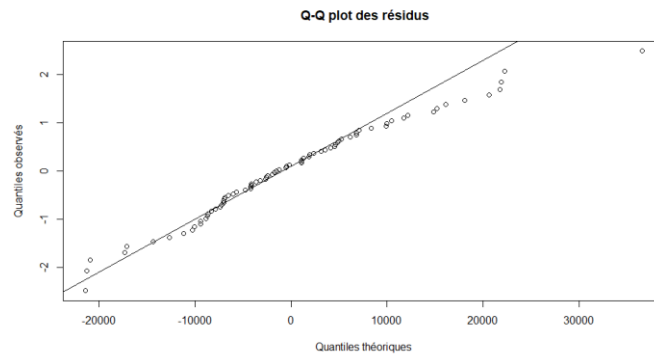


Annexe 31 : Evolution moyenne de la marge partielle (milliers d'euros/ha) prédit par le modèle 1

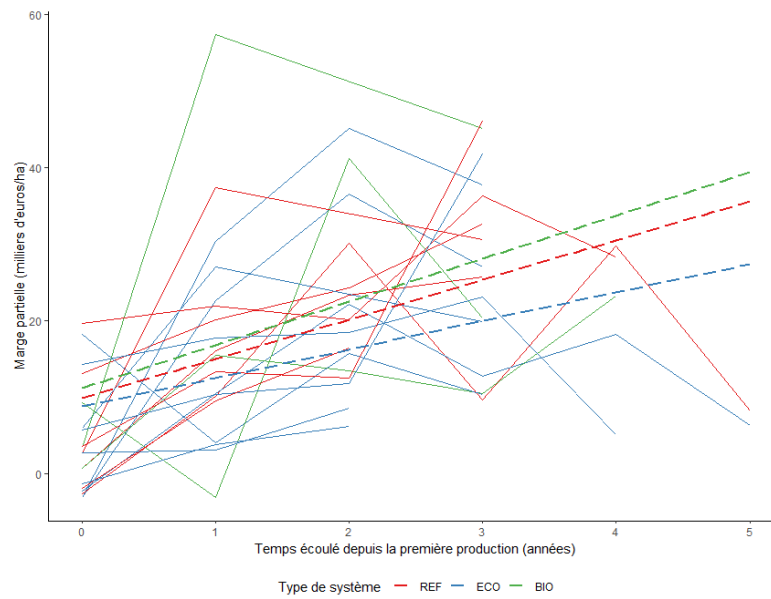


Annexe 32 : Analyse des résidus - Marge partielle- Modèle 1

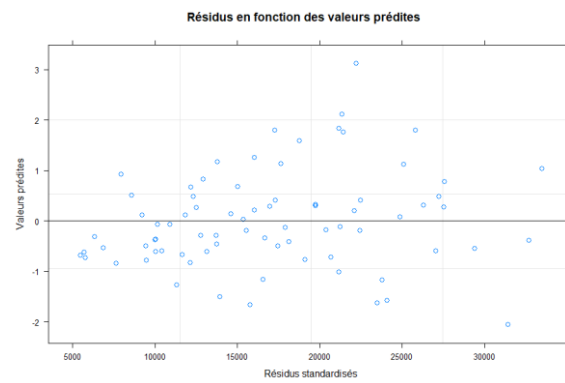


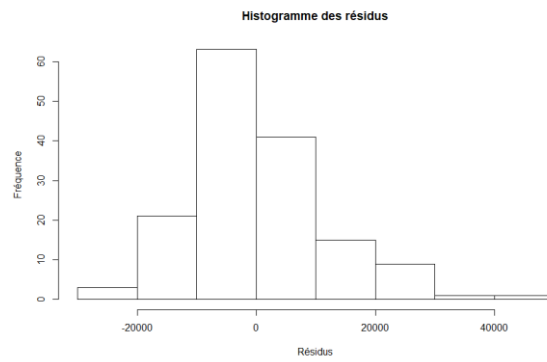
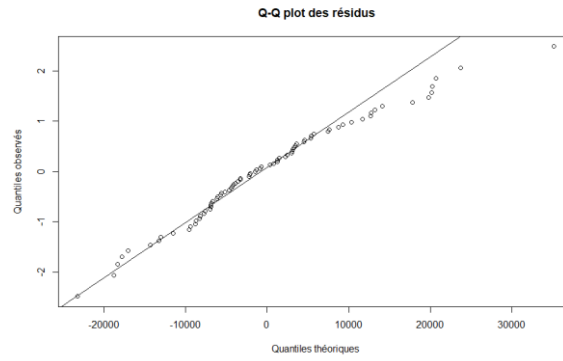


Annexe 33: Evolution moyenne de la marge partielle (milliers d'euros/ha) par type des systèmes prédite par le modèle 2

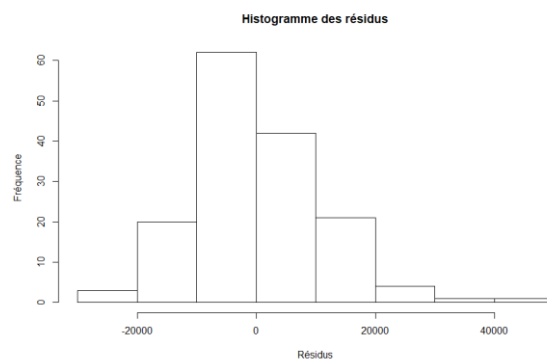
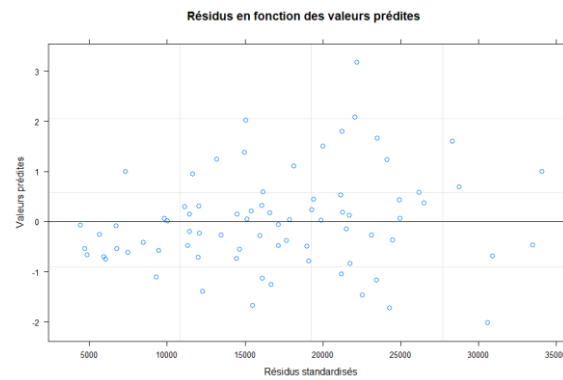


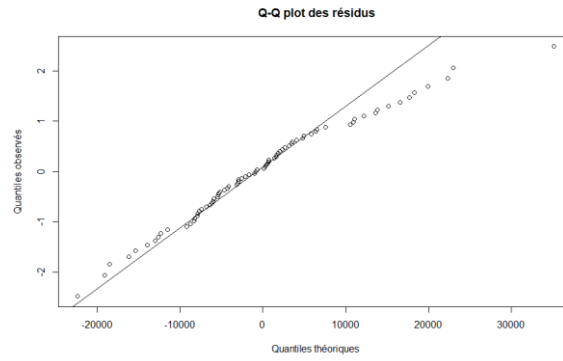
Annexe 34 : Analyse des résidus - Marge partielle - Modèle 2





Annexe 35 : Analyse des résidus - Marge partielle - Modèle 3





Annexe 36 : Estimation des effets fixes, des effets aléatoires et résultats du test du rapport de vraisemblance sur les modèles
- Modèle mixte sur la marge partielle

| M1 | | | | | M2 | | | M3 | | |
|---------------------|-----------|------------|-------|--------|------------|-------|-------|------------|--------|-------|
| Variables | Param. | Estimation | Std | P* | Estimation | Std | P* | Estimation | Std | P* |
| Intercept | β_0 | 9.549 | 2.368 | <0.001 | 9.917 | 3.909 | 0.014 | -0.059 | 10.062 | 0.995 |
| Temps | β_1 | 4.600 | 0.997 | <0.001 | 5.134 | 1.611 | 0.002 | 5.112 | 5.051 | 0.316 |
| Système (ECO) | β_2 | | | | -1.101 | 5.230 | 0.836 | 1.818 | 6.073 | 0.769 |
| Système (BIO) | β_3 | | | | 1.272 | 7.248 | 0.863 | 5.281 | 8.119 | 0.525 |
| Temps*système (ECO) | β_4 | | | | -1.413 | 2.203 | 0.524 | -1.164 | 2.894 | 0.689 |
| Temps*système (BIO) | β_5 | | | | 0.503 | 3.061 | 0.870 | 0.657 | 3.556 | 0.854 |
| IFT | β_6 | | | | | | | 0.412 | 0.384 | 0.288 |
| Temps*IFT | β_7 | | | | | | | -0.004 | 0.194 | 0.982 |

* Test de Wald

| M1 | | | M2 | | M3 | |
|--------------|------------|--------|------------|--------|------------|--------|
| Paramètre | Estimation | Std | Estimation | Std | Estimation | Std |
| σ_b^2 | 21.947 | 4.685 | 20.138 | 4.488 | 21.480 | 4.635 |
| σ_e^2 | 129.606 | 11.384 | 126.616 | 11.252 | 122.234 | 11.056 |

| Variables | M1 | M2 | M3 |
|---------------|--------|--------|--------|
| Intercept | <0.001 | <0.001 | <0.001 |
| Temps | <0.001 | <0.001 | <0.001 |
| Système | | 0.485 | 0.456 |
| Temps*système | | 0.739 | 0.808 |
| IFT | | | 0.189 |
| Temps*IFT | | | 0.839 |